

78041

STIC-Biotech/ChemLib

195022

M9

From: Bunner, Bridget  
Sent: Monday, July 10, 2006 10:14 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search request

Hi! I'd like to request a sequence search for case 10/724,806:

1. the nucleic acid sequence that encodes SEQ ID NO:4
2. the nucleic acid sequence of SEQ ID NO: 3

Thanks!

Bridget Bunner

Art Unit 1647  
Rem 4C65  
(571) 272-0881  
mailbox 4C70

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:29 ; Search time 9689 Seconds  
(without alignments)  
11503.804 Million cell updates/sec

Title: US-10-724-806-3  
Perfect score: 1743  
Sequence: 1 atgccttcacatgctggaag.....ctgaagataattacaatga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: gb\_env:\*  
2: gb\_pat:\*  
3: gb\_ph:\*  
4: gb\_pl:\*  
5: gb\_pr:\*  
6: gb\_ro:\*  
7: gb\_srs:\*  
8: gb\_sy:\*  
9: gb\_un:\*  
10: gb\_vl:\*  
11: gb\_ov:\*  
12: gb\_hcg:\*  
13: gb\_in:\*  
14: gb\_om:\*  
15: gb\_ba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	6	AF276872 Mus muscu
2	1731.8	99.4	4715	6	BC065089 Mus muscu
3	1730.2	99.3	1743	2	BD012720 High-affi
4	1730.2	99.3	1743	2	BD012720 High-affi
5	1730.2	99.3	1743	2	BD012720 High-affi
6	1723.8	98.9	1743	6	AX080443 Sequence
7	1616.6	92.7	1743	2	MMU01467 Mus muscu
8	1616.6	92.7	1743	2	BD012718 High-affi
9	1616.6	92.7	1743	2	BD012718 High-affi
10	1378.2	79.1	1743	2	AB030947 Rattus no
11	1375	78.9	1743	2	AR268949 Sequence
12	1375	78.9	1743	2	BD012719 High-affi
13	1375	78.9	1743	2	CQ722844 Sequence
14	1375	78.9	1743	5	E49871 High-affin
15	1375	78.9	1743	5	AF276871 Homo sapi
16	1375	78.9	1813	5	HS401466 Homo sapi
17	1375	78.9	2583	5	BC111525 Homo sapi
18	1297.4	74.4	2541	5	AB043997 Homo sapi
					BC111524 Homo sapi

19	1243.8	71.4	2112	14	BC112794 Bos tauru
20	841.4	48.3	2528	11	TM4420808 Torpedo m
21	699	40.1	1132	11	GA511267 Gallus ga
22	629.2	36.1	155210	6	AC102873 Mus muscu
23	629.2	36.1	257456	12	CT571242 Mus muscu
24	600.4	34.4	232792	12	AC106657 Rattus no
25	504.4	28.9	159918	12	AC182392 Pan trogl
26	502.8	28.8	2239	5	HS401466 Homo sapi
27	502.8	28.8	190043	5	AC009963 Homo sapi
28	437.8	25.1	3326	13	AY011119 Limulus p
29	421.2	24.2	3255	13	AY047521 Drosophila
30	372.6	21.4	2049	13	AY629593 Trichoplu
31	356.4	20.4	386	2	AX080449 Sequence
32	355.8	20.4	1731	2	BD012717 High-affi
33	355.8	20.4	1731	2	E49869 High-affi
34	355.8	20.4	1985	13	AB030946 Caenorhab
35	321	18.4	329	2	AX956120 Sequence
36	278.8	16.0	1729	2	CQ613836 Sequence
37	257.8	14.8	172639	12	CR925787 Danio rer
38	240.4	13.8	212798	11	BX901972 Zebrafish
39	237.6	13.6	1461	2	AX432086 Sequence
40	237.6	13.6	110000	15	AE017333_31 Continuation (32 o
41	237.6	13.6	110000	15	CF000002_31 Continuation (32 o
42	221.4	12.7	166000	11	AC174718 Gasterost
43	178	10.2	4223	2	CQ613835 Sequence
44	178	10.2	186989	13	AC007812 Drosophila
45	178	10.2	189116	13	AC009395 Drosophila

## ALIGNMENTS

RESULT 1  
AP276872 1743 bp mRNA linear ROD 28-FEB-2001  
LOCUS  
DEFINITION Mus musculus sodium and chloride-dependent high-affinity choline transporter mRNA, complete cds.  
ACCESSION AF276872  
VERSION AF276872.2 GI:13162669  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS 1 (bases 1 to 1743)  
TITLE Aparsundaram,S., Ferguson,S.M. and Blakely,R.D.  
JOURNAL Molecular cloning and characterization of human and murine high-affinity choline transporters  
Unpublished  
2 (bases 1 to 1743)  
AUTHORS Aparsundaram,S., Ferguson,S.M. and Blakely,R.D.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUN-2000) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA

REFERENCE  
AUTHORS 3 (bases 1 to 1743)  
TITLE Aparsundaram,S., Ferguson,S.M. and Blakely,R.D.  
JOURNAL Direct Submission  
JOURNAL Submitted (28-FEB-2001) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA  
REMARK  
COMMENT Sequence update by submitter  
On Feb 28 2001 this sequence version replaced gi:11527247.  
FEATURES  
Location/Qualifiers  
source  
1..1743  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..1743  
/codon\_start=1  
/product="sodium and chloride-dependent high-affinity choline transporter"

## CDS

1..1743  
/product="sodium and chloride-dependent high-affinity choline transporter"

/protein\_id="AA036945.2"  
/db\_xref="GI:13162670"  
/translation="MPFVEGLVALILFYLLIPLVGIWAAMTKNSGNPERSEALIV  
GRIQILVGEFTMTATFWGGYINGTGAIVPGCGGAMAHAPIGYLSILIGLGF  
AKPMRSKVTMTLDPKROIYGRMGGLIFALMGEMMAAIPALGATLSVITD  
VNISVYSAALAILLYLVLGGLYSVAATTVVQLFCIFLMLISVPALSHPATVIGFT  
AVHAYOSFWLITTESEVYTWLDFLLMLGGIPWQYFQRVSSSSATYAOVSLF  
AAGCLVWALPAICGAIASGTDMNQATVGPDPKTKEDMILFIVLQYLCPVYISF  
FGLGAVSAVMSADSILISASMEFARNIYQISFRONASDEKIVWAGIIVLVFASA  
TAMALLTKTVGLMTWLSLDLYIIFPOLLCVLFKGNITGVAVGIIFGLFRLTGG  
EPYLIQPLIFYPGYSDKNGINORPEPKTSMWTSPTNIVSGLAKYLPESGTLF  
PKLVEFADVVAHSENNDKTILVNENIKINELAPVAPROSLTISSTFTKEALDV  
DSFEGSGTEHNLQ"

## ORIGIN

Query Match 100.0%; Score 1743; DB 6; Length 1743;  
Best Local Similarity 100.0%; Pred. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTTCATGTTGGAAGAGCTGTAGCTATTTATCTTCTTACCTCTTATTTCTG 60  
1 ATGCTTTCATGTTGGAAGAGCTGTAGCTATTTATCTTCTTACCTCTTATTTCTG 60  
61 GTTGAATATGSGGCTGCATGGAACCAAAACACGCGCAACCCGAGAAGGCGAGTGA 120  
61 GTTGAATATGSGGCTGCATGGAACCAAAACACGCGCAACCCGAGAAGGCGAGTGA 120  
121 GCCATCATAGTCGGGGGCGTGACATTTGTTGTTGTTGTTGTTTAAACATGACAGCC 180  
121 GCCATCATAGTCGGGGGCGTGACATTTGTTGTTGTTGTTGTTTAAACATGACAGCC 180  
181 ACCTGGGTTGAGAGGCTACATCAATGAGGACAGCAAGAGAGCTGATGGGCGAGTTGT 240  
181 ACCTGGGTTGAGAGGCTACATCAATGAGGACAGCAAGAGAGCTGATGGGCGAGTTGT 240  
241 GGTCTAGCTGGGCTCATGCAACCATTTGATTTCTGAGCTCATTTTAAAGGCTGCG 300  
241 GGTCTAGCTGGGCTCATGCAACCATTTGATTTCTGAGCTCATTTTAAAGGCTGCG 300  
301 TTTTTCGGAACCTATGCGTTCGAAGGATATGACTATGTATGAGACCATTTCAACAG 360  
301 TTTTTCGGAACCTATGCGTTCGAAGGATATGACTATGTATGAGACCATTTCAACAG 360  
361 ATCTATGGAAGCGCATGGGCTGCTCTTCACTCTGCACTGATGAGGAGATGTTTC 420  
361 ATCTATGGAAGCGCATGGGCTGCTCTTCACTCTGCACTGATGAGGAGATGTTTC 420  
421 TGGGCTGAGCAATTTCTCTGCACTTAAAGGGGCAACATGAGGCTATCATTTGATGAT 480  
421 TGGGCTGAGCAATTTCTCTGCACTTAAAGGGGCAACATGAGGCTATCATTTGATGAT 480  
481 GTGAACATATCGGTATGTTCTCTGCACTCATTTGCATTTTAAACCTAGTGGGTGG 540  
481 GTGAACATATCGGTATGTTCTCTGCACTCATTTGCATTTTAAACCTAGTGGGTGG 540  
541 CTCTACTCTGTGGCATATGATGTTGTCTGCACTTCTGCACTTTTAAAGACCTGTGG 600  
541 CTCTACTCTGTGGCATATGATGTTGTCTGCACTTCTGCACTTTTAAAGACCTGTGG 600  
601 ATCACTGTCCTTTGGCCCTGTCACATCTGCACTGACGACATGGGATTTCAAGCTGG 660  
601 ATCACTGTCCTTTGGCCCTGTCACATCTGCACTGACGACATGGGATTTCAAGCTGG 660  
661 CATGCTAAATACCAAGTCCCTGGCTGGGCAACATTTGAATCAGTTGAAGTCAACCTGG 720  
661 CATGCTAAATACCAAGTCCCTGGCTGGGCAACATTTGAATCAGTTGAAGTCAACCTGG 720  
721 CTGGAATATTTCTGTATGATGCTGGGTGGAATCCATGGCAAGCTTATCCAGAG 780  
721 CTGGAATATTTCTGTATGATGCTGGGTGGAATCCATGGCAAGCTTATCCAGAG 780  
781 GTCCCTCTTCACTCAGCACTATGCTCAGGTATCTCTCTCTGCGAGCTTTTGGG 840  
781 GTCCCTCTTCACTCAGCACTATGCTCAGGTATCTCTCTCTGCGAGCTTTTGGG 840

781 GTCCCTCTTCACTCAGCACTATGCTCAGGTATCTCTCTCTGCGAGCTTTTGGG 840  
841 TGCCGTGATAGGCTCTTACCCGATATGATAGAGCTATTTGAGCTTCCAGACATGG 900  
841 TGCCGTGATAGGCTCTTACCCGATATGATAGAGCTATTTGAGCTTCCAGACATGG 900  
901 AACCAAGTCCCTACGGGTATCCAGATCCCAAGACTAAGAGGAAGCAGACATGTTTC 960  
901 AACCAAGTCCCTACGGGTATCCAGATCCCAAGACTAAGAGGAAGCAGACATGTTTC 960  
961 CCGATCTGTCGAGTACCTTGCCCTGTCATCTCTCTTGGGCTTGGCTGTT 1020  
961 CCGATCTGTCGAGTACCTTGCCCTGTCATCTCTCTTGGGCTTGGCTGTT 1020  
1021 TCAGCTGTCATAGTCTCAGCTGACTGTCATCTCTGCGGAGTTCTATGTTTGGT 1080  
1021 TCAGCTGTCATAGTCTCAGCTGACTGTCATCTCTGCGGAGTTCTATGTTTGGT 1080  
1081 CGGAATATCTACAGCTTCTCTTCCAGCAAAATGCATCAGCAAGAAATGTGGGCTC 1140  
1081 ATGAGGATCAGTGTGCTTGTGCGAGCATCTGCAACAGCCATGGCTTGTGCGAAG 1200  
1141 ATGAGGATCAGTGTGCTTGTGCGAGCATCTGCAACAGCCATGGCTTGTGCGAAG 1200  
1141 ATGAGGATCAGTGTGCTTGTGCGAGCATCTGCAACAGCCATGGCTTGTGCGAAG 1200  
1201 ACTGTGATGAGGCTCTGCTGACTGAGCTGAGCTTGTCTATCATCATCTTCCACAG 1260  
1201 ACTGTGATGAGGCTCTGCTGACTGAGCTGAGCTTGTCTATCATCATCTTCCACAG 1260  
1261 CTGCTCTGTATCTCTTATCAAGGAACCAACTTATGGGGCAGTGTGCTTATAT 1320  
1261 CTGCTCTGTATCTCTTATCAAGGAACCAACTTATGGGGCAGTGTGCTTATAT 1320  
1321 CTGCTCTGTATCTCTTATCAAGGAACCAACTTATGGGGCAGTGTGCTTATAT 1320  
1321 CTGCTCTGTATCTCTTATCAAGGAACCAACTTATGGGGCAGTGTGCTTATAT 1320  
1381 TTTTACCTGCTTATCTCTGACCAAGATGTATATCAATCAGAGTTCCCATTTAA 1440  
1381 TTTTACCTGCTTATCTCTGACCAAGATGTATATCAATCAGAGTTCCCATTTAA 1440  
1441 ACTCTCTCAGTGTACTCTTATCAAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
1441 ACTCTCTCAGTGTACTCTTATCAAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
1501 ACTCTCTCAGTGTACTCTTATCAAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
1501 ACTCTCTCAGTGTACTCTTATCAAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
1501 CTATTGGAAGTGAACCTTGCCCTGTCACAAATTTAGATGATTTGATGCTGTGTCAGG 1560  
1501 CTATTGGAAGTGAACCTTGCCCTGTCACAAATTTAGATGATTTGATGCTGTGTCAGG 1560  
1561 CACAGTGAAGAACATGAGCAACATCTTATGCTAGCAAGAAATGAAATTAATTAAT 1620  
1561 CACAGTGAAGAACATGAGCAACATCTTATGCTAGCAAGAAATGAAATTAATTAAT 1620  
1621 GAACCTTGAACCTGGAACCTGCGAGAGCTTAAACCTTCACTTCACTTCACTTAAAG 1680  
1621 GAACCTTGAACCTGGAACCTGCGAGAGCTTAAACCTTCACTTCACTTCACTTAAAG 1680  
1681 GAGGCTCTCTTATGATGATTTCCAGTCCGAGAGGGGTCTGGGACTGAAGATTAATTA 1740  
1681 GAGGCTCTCTTATGATGATTTCCAGTCCGAGAGGGGTCTGGGACTGAAGATTAATTA 1740  
1741 TGA 1743  
1741 TGA 1743  
1741 TGA 1743

RESULT 2  
BC065089 4715 bp mRNA linear ROD 30-JUN-2004  
LOCUS BC065089  
DEFINITION Mus musculinus solute carrier family 5 (cholesterol transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:8651099), complete cds.  
ACCESSION BC065089



KEYWORD	BC065089.1 GI:40787811
KEYWORD	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 4715)
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McEwan, K.J., Malek, J.A., Garabara, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W., Vallat, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Medan, A., Young, A.C., Shcherbenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherbenko, Y., Bouffard, G.G., Blakeley, R.W., Guttman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Richmond, J., Schmitt, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherzer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 4715)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: Dr. Jim Ij, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant. Web site: <a href="http://genome.uiowa.edu">http://genome.uiowa.edu</a> Contact: <a href="mailto:bento-soares@iowa.edu">bento-soares@iowa.edu</a> ; <a href="mailto:tom-casavant@iowa.edu">tom-casavant@iowa.edu</a> Bonaldo, M.F., Akabogu, I., Bair, T., Bair, T., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Schaefer, T., Smith, C., Shit, E., Tack, D., Trouk, K., Walters, J., Casavant, T., Soares, M.B.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: Place: Row: Column: 0 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11528509. Location/Qualifiers 1..4715 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="MGC:86023 IMAGE:6851099" /tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc" /clone_lib="NIH BMAP_Fyo" /lab_host="DH10B" /note="Vector: pYX-ASC" 1..4715 /gene="Slc5a7" /note="synonym: CHT1" /db_xref="GeneID:63993" /db_xref="MG1:1927126"
gene	

CDS		27..1769
		/gene="Slc5a7"
		/codon_start=1
		/product="solute carrier family 5 (chooline transporter), member 7"
		/protein_id="AA065089.1"
		/db_xref="GI:40787812"
		/db_xref="GeneID:63993"
		/db_xref="MG1:1927126"
		/translation="MSFHVGVLAIIILFYLLIFLVGVIAAKTKNSGNPERSAIVIGSDIGILVGGFTMTATWVGVGIVNGIAEYVYRPGCLAAQAPISGLISLIGLFFAKMRSGIVTMLDLPFOIIGKMGILLIFPALMGEMFAAFSAIGATISVIIDVUNISVIVSALIALIYTLVGLSVAYAVDVOLFCIFGLMI SVFALSHAVNDIGFTAVAKVSGPWLGTIESVEVYTWLDFLLMLGGLPMAVQORVYLSASSATYAQVLSFAAGGALVMAIPALCIGIIGASTPMNCAIYVGPCKEADMILPIYLQICLPEYISFAGVSAVAVMSSADSILASSPMANCAIYVLSPROASKEIIVWNRITVLVYGASATMALILKTYIYGMWYSSDLVYIIIPQLLCVAFIKTNNYVAVAGYIPGLFPIITGGEPFLYVQPLIFPYGYSIDKNGIYNORPEFKLTSVNSFFPNNICVSYAKLIFESGTLPEPLVDPAVVAARSEENMDKTIIVRNNIKLIELAVPKPROSLTSLSTFNKELLDVDDSPBEGSGTELDNQ"
ORIGIN		
Query Match	99.4%; Score 1731.8; DB 6; Length 4715;	
Best Local Similarity	99.4%; Pred. No. 0;	
Matches 1736; Conservative	0; Mismatches 7; Indels 0; Gaps 0;	
Oy	1	ATGCGCTTCCATGCGAGAAGACTGGTAGCTATATTCCTCTTCACTCCTATATTTCTG 60
Db	27	ATGCTTTCCACGTAGAGAAGACTGTAAGCTATATTCCTCTTCACTCCTATATTTCTG 86
Oy	61	GTTGGAAATATGCGCTGCATGGAAAAACCAAAAACAGCGGCAACCAAGAAAGCGCAGTGA 120
Db	87	GTTGGAAATATGCGCTGCATGGAAAAACCAAAAACAGCGGCAACCAAGAAAGCGCAGTGA 146
Oy	121	GCCATCATATGCGCGGCGCGCGCATGATGGTTGTTGGTGGTGGTTTATCCATGACAGCC 180
Db	147	GCCATCATATGCGCGGCGCGCGCATGATGGTTGTTGGTGGTGGTTTATCCATGACAGCC 206
Oy	181	ACCTGGGTTGAGAGGAGGCTACATCAATGGGACAGCAAGACGTATGGGCCAGGTTGT 240
Db	207	ACCTGGGTTGAGAGGAGGCTACATCAATGGGACAGCAAGACGTATGGGCCAGGTTGT 266
Oy	241	GCTTACCTGGGCTCATGACCCCATGGGATATCTCGAGTCTAAATTTTAGGGTGCTG 300
Db	267	GCTTACCTGGGCTCATGACCCCATGGGATATCTCGAGTCTAAATTTTAGGGTGCTG 326
Oy	301	TTTTTTCGAAACCTATAGCGTTCCAAAGGGAATGTGACTATGTAGACCCATTCAACAG 360
Db	327	TTTTTTCGAAACCTATAGCGTTCCAAAGGGAATGTGACTATGTAGACCCATTCAACAG 386
Oy	361	ATCTATGAAAGCGCATGGGTGGGCTGCTTCAATCCTTGCACTGATGGGAGAGATGTT 420
Db	387	ATCTATGAAAGCGCATGGGTGGGCTGCTTCAATCCTTGCACTGATGGGAGAGATGTT 446
Oy	421	TGGGCTGACGAATTTTCTCTGCATTTTGGGGCCACATCAGCGTATCATTTGATGTTGAT 480
Db	447	TGGGCTGACGAATTTTCTCTGCATTTTGGGGCCACATCAGCGTATCATTTGATGTTGAT 506
Oy	481	GTTGAACATATGCTATGTCTCTGCACTCATTTGCCATTCTTTATACCTTAGTGGGTGG 540
Db	507	GTTGAACATATGCTATGTCTCTGCACTCATTTGCCATTCTTTATACCTTAGTGGGTGG 566
Oy	541	CTTACCTGTGGGATATCATGATGTTGTCAGGTAATTCGCAATTTTATAGGACTGGG 600
Db	567	CTTACCTGTGGGATATCATGATGTTGTCAGGTAATTCGCAATTTTATAGGACTGGG 626
Oy	601	ATCAGTGTCCCTTTTGGCCCTGTCAATCTCTGACGTACCCGACATTCGATTCACAGCTGTG 660
Db	627	ATCAGTGTCCCTTTTGGCCCTGTCAATCTCTGACGTACCCGACATTCGATTCACAGCTGTG 686
Oy	661	CATGCTAAATACAGAGATCCCTGGCTGGGAACATTAATAGTTGAAGTTACACCTGG 720
Db	687	CATGCTAAATACAGAGATCCCTGGCTGGGAACATTAATAGTTGAAGTTACACCTGG 746

QY 721 CTGTATATATTTCTGTATATGTAGTGTGGTGAATCCATGCGCAAGCTTATCCAGAG 780  
DB 747 CTTGATATATTTCTGTATATGTAGTGTGGTGAATCCATGCGCAAGCTTATCCAGAG 806  
QY 781 GTCCCTCTTATCTCTGAGCCCATATGCTCAGTACTGTCTTCTCTGAGCTTTTGGG 840  
DB 807 GTCCCTCTTATCTCTGAGCCCATATGCTCAGTACTGTCTTCTCTGAGCTTTTGGG 866  
QY 841 TGCCCTGTATGAGTCTCTACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGATGG 900  
DB 867 TGCCCTGTATGAGTCTCTACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGATGG 926  
QY 901 AACCAAGCTGCTTACCGGGTATCCAGATCCCAAGACTTAAGAGAAACAGACATGATTC 960  
DB 927 AACCAAGCTGCTTACCGGGTATCCAGATCCCAAGACTTAAGAGAAACAGACATGATTC 986  
QY 961 CGGATGCTTGTGAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 987 CGGATGCTTGTGAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046  
QY 1021 TCAGCTGCTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 1047 TCAGCTGCTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106  
QY 1081 CGGAATATCTACAGCTTTCTTCAACAATAATGATCAGACAGAAATTTGTGTGCTC 1140  
DB 1107 CGGAATATCTACAGCTTTCTTCAACAATAATGATCAGACAGAAATTTGTGTGCTC 1166  
QY 1141 ATGAGATATCTACAGCTTTCTTCAACAATAATGATCAGACAGAAATTTGTGTGCTC 1200  
DB 1167 ATGAGATATCTACAGCTTTCTTCAACAATAATGATCAGACAGAAATTTGTGTGCTC 1226  
QY 1201 ACTGTGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
DB 1227 ACTGTGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286  
QY 1261 CTGCTCTGTGTATCTTCTCAAGAAACCAACTTATGSGGCGAGTGTCTGTATATTT 1320  
DB 1287 CTGCTCTGTGTATCTTCTCAAGAAACCAACTTATGSGGCGAGTGTCTGTATATTT 1346  
QY 1321 TTTGAGCTATCTCTGAGATTAATCTGAGAGAGAGCAATCTATCTGTGAGGCGCTTAATC 1380  
DB 1347 TTTGAGCTATCTCTGAGATTAATCTGAGAGAGAGCAATCTATCTGTGAGGCGCTTAATC 1406  
QY 1381 TTTCTACCTGCTTATTAATCTGAGCAAGATGATATATCAATCAGAGGTTCCATTTAA 1440  
DB 1407 TTTCTACCTGCTTATTAATCTGAGCAAGATGATATATCAATCAGAGGTTCCATTTAA 1466  
QY 1441 ACTCTCTCAATGCTTATCTCAATCTTCAATCAATTTGTGTCTTATCTAGCCAAATAT 1500  
DB 1467 ACTCTCTCAATGCTTATCTCAATCTTCAATCAATTTGTGTCTTATCTAGCCAAATAT 1526  
QY 1501 CTATTTGAAAGTGAACCTTGTCTCAAAATAGATTTGATGAGTGTGTGTGTGTGTGTGT 1560  
DB 1527 CTATTTGAAAGTGAACCTTGTCTCAAAATAGATTTGATGAGTGTGTGTGTGTGTGTGT 1586  
QY 1561 CACAGTGAAGAGACATGAGCAAGACCAATTTAGTCAGAAATGAATAATCAATTTAAAT 1620  
DB 1587 CACAGTGAAGAGACATGAGCAAGACCAATTTAGTCAGAAATGAATAATCAATTTAAAT 1646  
QY 1621 GAACTTGAACCTTGT 1680  
DB 1647 GAACTTGAACCTTGT 1706  
QY 1681 GAGGCGCTCTTGT 1740  
DB 1707 GAGGCGCTCTTGT 1766  
QY 1741 TGA 1743  
DB 1767 TGA 1769

RESULT 3  
BD012720 1743 bp DNA linear PAT 02-AUG-2002  
LOCUS  
DEFINITION High-affinity choline transporter.  
ACCESSION BD012720.1 GI:22092909  
VERSION WO 0116315-A/4.  
KEYWORDS Mus musculus (house mouse)  
SOURCE  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Mus.  
REFERENCE  
Haga, T. and Okuda, T.  
TITLE High-affinity choline transporter  
JOURNAL Patent: WO 0116315-A 08-MAR-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA  
COMMENT  
OS Mus musculus (mouse)  
PN WO 0116315-A/4  
PD 08-MAR-2001  
PF 18-AUG-2000 WO 2000JP005545  
PR 27-AUG-1999 JP 99P 240642.27-DEC-1999 JP 99P 368991 PI  
TATSUYA HAGA, TAKASHI OKUDA  
PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC  
A61K38/17,  
A61K45/00, A61P25/28, G01N33/53, A01K67/027  
CC  
FH Key Location/Qualifiers  
FT CDS (1)..(1743).  
FEATURES  
source  
ORIGIN  
Query Match 99.3%; Score 1730.2; DB 2; Length 1743;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 ATGCTTTCCATGATGAGAGCTGTGATCTATATCTTCTCTCTCTCTCTCTCTCTCTCTCT 60  
DB 1 ATGCTTTCCATGATGAGAGCTGTGATCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
QY 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGCAACCCAGAGAGCGAGTGA 120  
DB 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGCAACCCAGAGAGCGAGTGA 120  
QY 121 GCCATCATAGTGGGGGCGCGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180  
DB 121 GCCATCATAGTGGGGGCGCGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180  
QY 121 ACCGTGGTGGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 121 ACCGTGGTGGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 241 GGTCTAGCTGGGCTCATGACCCATTTGATATTTCTGAGTCTAATTTAGTGTGTG 300  
DB 241 GGTCTAGCTGGGCTCATGACCCATTTGATATTTCTGAGTCTAATTTAGTGTGTG 300  
QY 301 TTTTGTGGAAACCTATGCTTCCAGAGATATGATATGATTTGATTTGATTTGATTTGATTT 360  
DB 301 TTTTGTGGAAACCTATGCTTCCAGAGATATGATATGATTTGATTTGATTTGATTTGATTT 360  
QY 361 ATCTATGAAAGCGCATGAGTGGGCTGCTCTTCAATCCCTGCACTATGGAGAGATGTTT 420  
DB 361 ATCTATGAAAGCGCATGAGTGGGCTGCTCTTCAATCCCTGCACTATGGAGAGATGTTT 420  
QY 421 TGGGCTGAGCAATTTTCTGCTATGAGGAGCCACATCAGCGTATCATTTGATTTGATTT 480  
DB 421 TGGGCTGAGCAATTTTCTGCTATGAGGAGCCACATCAGCGTATCATTTGATTTGATTT 480

QY 481 GTGAACATATCGGTCAATGCTCTGCACTATTCGACATCTTATATCCCTAGTGGTGG 540  
 DB 481 GTGAACATATCGGTCAATGCTCTGCACTATTCGACATCTTATATCCCTAGTGGTGG 540  
 QY 541 CTCTACTCTGTGGCATATATCTGATGTTGTCACGCTATTTGCACTTTTATAGGACTGTG 600  
 DB 541 CTCTACTCTGTGGCATATATCTGATGTTGTCACGCTATTTGCACTTTTATAGGACTGTG 600  
 QY 601 ATCAGAGTCCCTTTGGCCGTGCATATCTGCACTGACGACATCGGATTCACAGCTGTG 660  
 DB 601 ATCAGAGTCCCTTTGGCCGTGCATATCTGCACTGACGACATCGGATTCACAGCTGTG 660  
 QY 661 CATGCTAAATACAGAGTCCCTGGCTGGGAAACATTGAATCAGTTGAAGCTACACTGG 720  
 DB 661 CATGCTAAATACAGAGTCCCTGGCTGGGAAACATTGAATCAGTTGAAGCTACACTGG 720  
 QY 721 CTTGATATATTTCTGTTATATGATGCTGGTGGATCCCATGGCAAGCTTCCAGAGG 780  
 DB 721 CTTGATATATTTCTGTTATATGATGCTGGTGGATCCCATGGCAAGCTTCCAGAGG 780  
 QY 781 GTTCCTCTTCATCTCAGGCACTATGCTCAGGTAAGTCTTCCGGCAGCTTTGGG 840  
 DB 781 GTTCCTCTTCATCTCAGGCACTATGCTCAGGTAAGTCTTCCGGCAGCTTTGGG 840  
 QY 841 TGCCTGTGATGCTCTACCCGCAATATGATAGAGCTATTTGAGCTTCCAGAGCTGG 900  
 DB 841 TGCCTGTGATGCTCTACCCGCAATATGATAGAGCTATTTGAGCTTCCAGAGCTGG 900  
 QY 901 AACCCAGACTGCTACGGGTATCCAGATCCCAAGACTTAAAGGAGAGACAGATGATTTCTC 960  
 DB 901 AACCCAGACTGCTACGGGTATCCAGATCCCAAGACTTAAAGGAGAGACAGATGATTTCTC 960  
 QY 961 CGGATGCTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 DB 961 CGGATGCTCTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 QY 1021 TCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 DB 1021 TCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 QY 1081 CGGAATATCTACGAGCTTCTTCCAGCAAAATGCAATGCAAGCAAGAAATTTGTGGGCTC 1140  
 DB 1081 CGGAATATCTACGAGCTTCTTCCAGCAAAATGCAATGCAAGCAAGAAATTTGTGGGCTC 1140  
 QY 1141 ATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 DB 1141 ATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 QY 1201 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 DB 1201 ACTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 1261 CTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 DB 1261 CTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
 QY 1321 TTGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 DB 1321 TTGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 QY 1381 TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 DB 1381 TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
 QY 1441 ACTGCT 1500  
 DB 1441 ACTGCT 1500  
 QY 1501 CTATTTGAAAGTGAAGCTTGGCTCCAAATTTAGATGATTTGATGCTGCTGCTGCTGCTGCTGCT 1560  
 DB 1501 CTATTTGAAAGTGAAGCTTGGCTCCAAATTTAGATGATTTGATGCTGCTGCTGCTGCTGCTGCT 1560  
 QY 1561 CACAGTGAAGAGAACATGCAAGAACCATTTCTAGTCAAGAAATGCAAAATTTAAAT 1620

DB 1561 CACAGTGAAGAGAACATGCAAGAACCATTTCTAGTCAAGAAATGCAAAATTTAAAT 1620  
 QY 1621 GAACTTGCACTGCTGTAAGAACTGGGAGAGCTTAACTTCACTTCACTTCACTTCACTTCACTTCACT 1680  
 DB 1621 GAACTTGCACTGCTGTAAGAACTGGGAGAGCTTAACTTCACTTCACTTCACTTCACTTCACTTCACT 1680  
 QY 1681 GAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 DB 1681 GAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 QY 1741 TGA 1743  
 DB 1741 TGA 1743  
 RESULT 4  
 E49872 1743 bp DNA linear PAT 27-AUG-2002  
 LOCUS High-affinity choline transporter.  
 DEFINITION High-affinity choline transporter.  
 E49872 1 GI:22554903  
 ACCESSION E49872.1  
 VERSION E49872.1  
 KEYWORDS JP 2001136976-A/4.  
 SOURCE Mus sp.  
 ORGANISM Mus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1743)  
 AUTHORS Haga,T. and Okuda,T.  
 TITLE High-affinity choline transporter  
 JOURNAL Patent: JP 2001136976-A 4 22-MAY-2001;  
 SCIENCE & TECH AGENCY  
 COMMENT OS Mus sp. (mouse)  
 PN JP 2001136976-A/4  
 PD 22-MAY-2001  
 PF 27-DEC-1999 JP 1999368991  
 PI TATSUYA HAGA,TAKASHI OKUDA  
 PC C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,  
 PC C12N5/10  
 PC C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC  
 FH Key Location/Qualifiers  
 FT CDS Location/Qualifiers  
 1..1743  
 /organism="Mus sp."  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10095"  
 ORIGIN  
 Query Match 99.3%; Score 1730.2; DB 2; Length 1743;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 ATGCTTTTCATGTGAGAGAGCTGAGCTATATCTCTTACCTCTTATATTTCTG 60  
 DB 1 ATGCTTTTCATGTGAGAGAGCTGAGCTATATCTCTTACCTCTTATATTTCTG 60  
 QY 61 GTTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 61 GTTGAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 QY 121 GCATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 121 GCATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 181 ACTGCT 240  
 DB 181 ACTGCT 240  
 QY 241 GGTCTAGCTTGGGCTGATGACCATTCATTTCTGAGTCTAAATTTAGGTGCTG 300  
 DB 241 GGTCTAGCTTGGGCTGATGACCATTCATTTCTGAGTCTAAATTTAGGTGCTG 300

QY 301 TTTTGGGAAACCTATGCGTTCCAAAGGATATGACTATGTAGACCATTTCAACAG 360  
DB 301 TTTTGGGAAACCTATGCGTTCCAAAGGATATGACTATGTAGACCATTTCAACAG 360  
QY 361 ATCTATGAAAGCGATGGGTGGGCTCTTCACTCCCTGCACTGATGGGAAAGATGTTTC 420  
DB 361 ATCTATGAAAGCGATGGGTGGGCTCTTCACTCCCTGCACTGATGGGAAAGATGTTTC 420  
QY 421 TGGGTGAGCAATTTTCTGTCATTTAGGGGCGCACATCAGCGTATCTATGATGAT 480  
DB 421 TGGGTGAGCAATTTTCTGTCATTTAGGGGCGCACATCAGCGTATCTATGATGAT 480  
QY 481 GTGAACATATGGCTATGTTCTCTGCACTCATTTGCCATTTCTTAACTTACCTAGTGGGTGG 540  
DB 481 GTGAACATATGGCTATGTTCTCTGCACTCATTTGCCATTTCTTAACTTACCTAGTGGGTGG 540  
QY 541 CTCTACTGAGGCGATATGATGTTGTCAGCTATTCGATTTTATAGGACTGG 600  
DB 541 CTCTACTGAGGCGATATGATGTTGTCAGCTATTCGATTTTATAGGACTGG 600  
QY 601 ATCACTGTCCTTTTGGCTTGCATCTGTCAGTCAACGACATGCGATTTCACTGTTG 660  
DB 601 ATCACTGTCCTTTTGGCTTGCATCTGTCAGTCAACGACATGCGATTTCACTGTTG 660  
QY 661 CATGCTAAATACAGAGTCCCTGGCTGGGACCATTTGAATCATGTTGAAGTCTACCTGG 720  
DB 661 CATGCTAAATACAGAGTCCCTGGCTGGGACCATTTGAATCATGTTGAAGTCTACCTGG 720  
QY 721 CTGTAATATTTCTGTTATGATGTCGGGTGGATCCCATGAGGCGTCTTCCAGAG 780  
DB 721 CTGTAATATTTCTGTTATGATGTCGGGTGGATCCCATGAGGCGTCTTCCAGAG 780  
QY 781 GTCTCTCTTCACTTCAGCACTATGCTCAGGATCTGTTCTTCTGCGACCTTTTGGG 840  
DB 781 GTCTCTCTTCACTTCAGCACTATGCTCAGGATCTGTTCTTCTGCGACCTTTTGGG 840  
QY 841 TGCCGTGATGAGCTCTTCAACCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900  
DB 841 TGCCGTGATGAGCTCTTCAACCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900  
QY 901 AACGAGATGCGCTTACGGGATCCAGATCCCAAGACTAAGGAGGAGGACATGATTC 960  
DB 901 AACGAGATGCGCTTACGGGATCCAGATCCCAAGACTAAGGAGGAGGACATGATTC 960  
QY 961 CCGATCGTTTGCAGTACCTTGCCTGTGATCATCTCTTCTTGGGCTTGGTCTGTT 1020  
DB 961 CCGATCGTTTGCAGTACCTTGCCTGTGATCATCTCTTCTTGGGCTTGGTCTGTT 1020  
QY 1021 TCAGTGTGTCATATCTCTCACTGATCTGTCATCTGTCGGGAGTTCTATGTTGCT 1080  
DB 1021 TCAGTGTGTCATATCTCTCACTGATCTGTCATCTGTCGGGAGTTCTATGTTGCT 1080  
QY 1081 CGGAATATCTACACACTTCTTCCAGACAAATGATCAGACAAAGAAATTTGTGGGTC 1140  
DB 1081 CGGAATATCTACACACTTCTTCCAGACAAATGATCAGACAAAGAAATTTGTGGGTC 1140  
QY 1141 ATGAGGATCACTGCTGTGTGTTGCGAGCATCTGCAACAGCCATGCTTGTGTCAGAG 1200  
DB 1141 ATGAGGATCACTGCTGTGTGTTGCGAGCATCTGCAACAGCCATGCTTGTGTCAGAG 1200  
QY 1201 ACTGTGTATGGGCTCTGCTGACTGAGCTCTGACCTTGTCTATCATCATCTTCCACAG 1260  
DB 1201 ACTGTGTATGGGCTCTGCTGACTGAGCTCTGACCTTGTCTATCATCATCTTCCACAG 1260  
QY 1261 CTGCTCTGTATCTTCTTCACTCAAGAAACCAACTTATGAGGCGAGTTCTGTTATAT 1320  
DB 1261 CTGCTCTGTATCTTCTTCACTCAAGAAACCAACTTATGAGGCGAGTTCTGTTATAT 1320  
QY 1321 TTTGACATATTCCTGAGAAATTACTGAGAGAGCCATATCTATCTTGCAGCCCTTAATC 1380  
DB 1321 TTTGACATATTCCTGAGAAATTACTGAGAGAGCCATATCTATCTTGCAGCCCTTAATC 1380

QY 1381 TTCTACCTGCTTATTTACTCTGACAGAGATGATATATCAATCAGAGGTTCCATTAA 1440  
DB 1381 TTCTACCTGCTTATTTACTCTGACAGAGATGATATATCAATCAGAGGTTCCATTAA 1440  
QY 1441 ACTCTCTCAGTATACCTCATCTTCTTCAACAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
DB 1441 ACTCTCTCAGTATACCTCATCTTCTTCAACAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
QY 1501 CTATTTGAAAGTGAACCTTCCCTCCAAATTTAGATGATTTGATGCTGTTGTGCGAAG 1560  
DB 1501 CTATTTGAAAGTGAACCTTCCCTCCAAATTTAGATGATTTGATGCTGTTGTGCGAAG 1560  
QY 1561 CACAGTGAAGAGACATGAGACCAACCTTCTAGTCAGAAATGAAATATCAATTAAT 1620  
DB 1561 CACAGTGAAGAGACATGAGACCAACCTTCTAGTCAGAAATGAAATATCAATTAAT 1620  
QY 1621 GAACCTTGACCTGGAACCTCCGAGAGCCCTAACCTCAGTTCAACTTCAACCAATAG 1680  
DB 1621 GAACCTTGACCTGGAACCTCCGAGAGCCCTAACCTCAGTTCAACTTCAACCAATAG 1680  
QY 1681 GAGGCTCTCTGATGTTGATTTCCAGTCCGAGAGGCTCTGGGACTGAATATTTACA 1740  
DB 1681 GAGGCTCTCTGATGTTGATTTCCAGTCCGAGAGGCTCTGGGACTGAATATTTACA 1740  
QY 1741 TGA 1743  
DB 1741 TGA 1743

RESULT 5  
AX080443 4938 bp DNA linear PAT 22-FEB-2001  
LOCUS AX080443  
DEFINITION Sequence 1 from Patent WO0078950.  
ACCESSION AX080443  
VERSION AX080443.1 GI:13159872  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 Siersega, M. and Albrandt, K.  
AUTHORS Differentially expressed genes in the adipocytes of obese mice  
TITLE JOURNAL Patent: WO 0078950-A 1 28-DEC-2000;  
AMYLIN PHARMACEUTICALS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..4938  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"  
/note="P4P6B1"

ORIGIN  
Query Match 99.3%; Score 1730.2; DB 2; Length 4938;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 ATGCTTTCCATGTTGGAAGACCTGTTAGCTATTTACTCTTCTCACTCTTATATTTCTG 60  
DB 247 ATGCTTTCCACGTAGGAAGACTGTTAGCTATTTACTCTTCTCACTCTTATACTTCTG 306  
QY 61 GTTGAATATGAGGCTGTCATGGAACCAAAACAGCGGCAACCCAGAGAGCGCAGTGA 120  
DB 307 GTTGAATATGAGGCTGTCATGGAACCAAAACAGCGGCAACCCAGAGAGCGCAGTGA 366  
QY 121 GGCATCATATGTCGGGGCGCGTGAACATTTGTTGTTGTTGTTGTTTACATGACAGCC 180  
DB 367 GGCATCATATGTCGGGGCGCGTGAACATTTGTTGTTGTTGTTTACATGACAGCC 426  
QY 181 ACTTGGTTTGAAGAGGCTATCATCAATGGAACAGCAGAGAGAGTATGAGGCGCAGTTGT 240  
DB 427 ACTTGGTTTGAAGAGGCTATCATCAATGGAACAGCAGAGAGAGTATGAGGCGCAGTTGT 486

QY 241 GGCTAGCTTGGGCTCATGCAACCATTGATATCTCTGAGTCTAATTTAGGTGCTG 300  
 Db 487 GGCTTAGCTTGGGCTCAGGCAACCATTTGATATCTCTGAGTCTAATTTAGGTGCTG 546  
 QY 301 TTTTTCGCAAACTTATGGGTTCAAGGATATGTGACTATGTTAAGCCATTCAACG 360  
 Db 547 TTTTTCGCAAACTTATGGGTTCAAGGATATGTGACTATGTTAAGCCATTCAACG 606  
 QY 361 ATCTATGAAAGGCAATGGGTTGGGCTCTTTCATCCCTGCACTGATGGGAGATGTC 420  
 Db 607 ATCTATGAAAGGCAATGGGTTGGGCTCTTTCATCCCTGCACTGATGGGAGATGTC 666  
 QY 421 TGGGCTGCAACAATTTCTCTGCAATTAAGGAGCCACCATGAGCTGATTCATTTGAT 480  
 Db 667 TGGGCTGCAACAATTTCTCTGCAATTAAGGAGCCACCATGAGCTGATTCATTTGAT 726  
 QY 481 GTGAACATATCGGCTATTTGTTCTGCACTTCATTTGCCATTTCTTATTAACCTAGTGG 540  
 Db 727 GTGAACATATCGGCTATTTGTTCTGCACTTCATTTGCCATTTCTTATTAACCTAGTGG 786  
 QY 541 CTCTACTCTGAGCAATATCTGATGTTGTCAGCACTATTCGCACTTTTATAGACTGTG 600  
 Db 787 CTCTACTCTGAGCAATATCTGATGTTGTCAGCACTATTTCTGCACTTTTATAGACTGTG 846  
 QY 601 ATCAGTGTCCCTTTTGGCTGTCACATCCCTGAGTCAACGACATCGGATTCACAGCTGTG 660  
 Db 847 ATCAGTGTCCCTTTTGGCTGTCACATCCCTGAGTCAACGACATCGGATTCACAGCTGTG 906  
 QY 661 CATGCTAAATACCAGAGTCCCTGGCTGGGAAACATTGAATCAATTGAAGTCTACACTGTG 720  
 Db 907 CATGCTAAATACCAGAGTCCCTGGCTGGGAAACATTGAATCAATTGAAGTCTACACTGTG 966  
 QY 721 CTTGATATTTTCTGTTATTTGATGCTGGGTTGAATCCCAAGCACTTCCAGAG 780  
 Db 967 CTTGATATTTTCTGTTATTTGATGCTGGGTTGAATCCCAAGCACTTCCAGAG 1026  
 QY 781 GTCTCTCTTTCATCTCTCAGCAACCTATGCTCAGAGTCTGCTCTTCTGGAGCTTTGGG 840  
 Db 1027 GTCTCTCTTTCATCTCTCAGCAACCTATGCTCAGAGTCTGCTCTTCTGGAGCTTTGGG 1086  
 QY 841 TGCCTGTGATGCTCTACCCGCAATATGATGAGCTATTTGAGCTTTCACAGACTGG 900  
 Db 1087 TGCCTGTGATGCTCTACCCGCAATATGATGAGCTATTTGAGCTTTCACAGACTGG 1146  
 QY 901 AACCGAATCGCTCAACGGTATCCAGATCCCAAGAACTTAAGAGAACACATGATTTCTC 960  
 Db 1147 AACCGAATCGCTCAACGGTATCCAGATCCCAAGAACTTAAGAGAACACATGATTTCTC 1206  
 QY 961 CCGATGCTCTGAGTACCTCTGCTGATCATCTCTTTGGGCTTGGTGTGCTT 1020  
 Db 1207 CCGATGCTCTGAGTACCTCTGCTGATCATCTCTTTGGGCTTGGTGTGCTT 1266  
 QY 1021 TGAAGCTGCTGATGCTCTGAGTGAATGCTGTCATCTGTGGCGAGTTCTATGTTTGGCT 1080  
 Db 1267 TGAAGCTGCTGATGCTCTGAGTGAATGCTGTCATCTGTGGCGAGTTCTATGTTTGGCT 1326  
 QY 1081 CGGAATATCTACAGCTTCTTCCAGAAAATGCACTGACAAAGAAATGTGTGGTTC 1140  
 Db 1327 CGGAATATCTACAGCTTCTTCCAGAAAATGCACTGACAAAGAAATGTGTGGTTC 1386  
 QY 1141 ATGAGGATCACTGCTGCTGTTGGAGAGATCTGCAACAGCCATGCTTGTGCTGAGAG 1200  
 Db 1387 ATGAGGATCACTGCTGCTGTTGGAGAGATCTGCAACAGCCATGCTTGTGCTGAGAG 1446  
 QY 1201 ACTGTGTATGGGCTCTGATACCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260  
 Db 1447 ACTGTGTATGGGCTCTGATACCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1506  
 QY 1261 CTGCTGTGTATCTCTTCAATCAAGAGAACCAACTTATGGGCACTGCTGTGTTATTT 1320  
 Db 1507 CTGCTGTGTATCTCTTCAATCAAGAGAACCAACTTATGGGCACTGCTGTGTTATTT 1566

QY 1321 TTGGACTATTCCTGAGAAATTACTGAGAGAGACCATATCTATACTTGAGCCCTTAATC 1380  
 Db 1567 TTGGACTATTCCTGAGAAATTACTGAGAGAGACCATATCTATACTTGAGCCCTTAATC 1626  
 QY 1381 TTCTACCTCGTTTATTAATCTGACAAAGATGATATATCAATAGAGTTCCATTAA 1440  
 Db 1627 TTCTACCTCGTTTATTAATCTGACAAAGATGATATATCAATAGAGTTCCATTAA 1686  
 QY 1441 ACTCTCCATGATGTTACTCATCTTTTACCAATTTGTTCTTATCTAGCCAGATAT 1500  
 Db 1687 ACTCTCCATGATGTTACTCATCTTTTACCAATTTGTTCTTATCTAGCCAGATAT 1746  
 QY 1501 CTATTTGAAAGTGAACCTTGCTCCAAAATTAAGATGATTTGATGCTGTGCGAAG 1560  
 Db 1747 CTATTTGAAAGTGAACCTTGCTCCAAAATTAAGATGATTTGATGCTGTGCGAAG 1806  
 QY 1561 CACAGTGAAGAGAACCTGACAAAGACATTTCTAGTACAAATGAAATATCAATTAAT 1620  
 Db 1807 CACAGTGAAGAGAACCTGACAAAGACATTTCTAGTACAAATGAAATATCAATTAAT 1866  
 QY 1621 GAACTTGACCTGTGAAACCTGGGAGAGCTTAACCTCAGTTCAACTTACCAATTAAG 1680  
 Db 1867 GAACTTGACCTGTGAAACCTGGGAGAGCTTAACCTCAGTTCAACTTACCAATTAAG 1926  
 QY 1681 GAGGCCCTCTGATGTTGATTCAGTCCGAGAGGCTCTGGACTGAAGTAAATTTACA 1740  
 Db 1927 GAGGCCCTCTGATGTTGATTCAGTCCGAGAGGCTCTGGACTGAAGTAAATTTACA 1986  
 QY 1741 TGA 1743  
 Db 1987 TGA 1989

RESULT 6  
 MMU401467 1743 bp mRNA linear ROD 16-AUG-2000  
 LOCUS  
 DEFINITION  
 ACCESSION  
 AJ401467  
 VERSION  
 AJ401467.1 GI:9843808  
 KEYWORDS  
 CHT1 gene; high affinity choline transporter.  
 SOURCE  
 Mus musculus  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.  
 1  
 Wieland, A., Bonisch, H. and Bruns, M.  
 Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure  
 Unpublished  
 2 (bases 1 to 1743)  
 Bruns, M.  
 Direct Submission  
 Submitted (14-AUG-2000) Bruns M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
 FEATURES  
 source  
 1..1743  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/cJ"  
 /db\_xref="taxon:10090"  
 /tissue\_type="brainstem"  
 1..1743  
 /gene="CHT1"  
 1..1743  
 /gene="CHT1"  
 /function="sodium- and chloride-dependent reuptake of choline"  
 /experiment="experimental evidence, no additional details recorded"  
 /codon\_start=1  
 /product="high affinity choline transporter"

/protein\_id="CAC03719.1"  
/db\_xref="GI:9843809"  
/db\_xref="GOA:Q9BSM5"  
/db\_xref="UniProtKB/TrEMBL:Q9BSM5"  
/translation="MSFHEGLVATILFLILFLVGLVMAWTKNSGNDPEHSEALIV  
GARDIQLVGGTMTATVWGGVINGTVAAYVPGGLAMANOAPKYSLSLIGLGLP  
AKPMRSKYVTMLDPFOQYGRMGSLFLIPALMBEMTMAALFSLGHTISVITDID  
VNIISVIALIAILTYLVGLYSVAIVDVLCIFIGIMISVPALSHPATDIDGT  
AVHAKYQSPWLCTIESVEVYTWLDFLMLAGIPMOAYFORVLSSSSATVAQVLSYL  
AAGCLVMAIPALICIGALGASIDMNQATYVDPKTKEDMILPVLQYLPVYISF  
FGLGASAAVMSADSSILASMSFARNYOLSFROMASDEKILVWMTITVVFASAT  
TAMALLTKTVGLWYLSIDLYIILFPOLICVLFIKGTNYVAAGVYIEGLFLRTG  
EPLYIOLPIEPLYGYSDKNGIYNORPEKTSMTSPPTNI CVSTYLAKYLFESSTLP  
PKLDFVDAVVARHSEBNMDKITLVARNENIKNELAPVAKRQSLTUSSTFTKEALLDV  
DSSPESGTEEDMLQ"

## ORIGIN

Query Match 98.9%; Score 1723.8; DB 6; Length 1743;

Beet Local Similarity 99.3%; Pred. No. 0; Mismatches 12; Indels 0; Gaps 0;

Matches 1731; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

1 ATGCTTTCATGTGGAGAGCTGTAGCTATTATCTCTTACCTCTTAATTTCG 60  
1 ATGCTTTCAGTGAAGAGCTGTAGCTATTATCTCTTACCTCTTAATTTCG 60  
61 GTTGGAAATATGGGCTGCATGAAAAACAGCGGCAACCCAGAGAGCGAGTAA 120  
61 GTTGGAAATATGGGCTGCATGAAAAACAGCGGCAACCCAGAGAGCGAGTAA 120  
121 GCCATCATVATGCGGGGCGGTGACATVGTGTTGTTGTTGTTGTTTTCATGACAGCC 180  
121 GCCATCATVATGCGGGGCGGTGACATVGTGTTGTTGTTGTTTTCATGACAGCC 180  
181 ACCTGGGTTGAGAGAGCTTACATCAATGGAGACAGAAAGAGTATGGGCGAGTTGT 240  
181 ACCTGGGTTGAGAGAGCTTACATCAATGGAGACAGAAAGAGTATGGGCGAGTTGT 240  
241 GGTCTAGCTTGGGCTCATGACACCATTTGATATCTCTAGCTTAATTTTAAAGTGTCTG 300  
241 GGTCTAGCTTGGGCTCATGACACCATTTGATATCTCTAGCTTAATTTTAAAGTGTCTG 300  
301 TTTTTCGGAACCTATGCGGTTCCAAAGGATATGTACTATGTATGACCATTTTCAACAG 360  
301 TTTTTCGGAACCTATGCGGTTCCAAAGGATATGTACTATGTATGACCATTTTCAACAG 360  
361 ATCTATGAAAAGCGATGGGTGGGCTGCTTTCATCCCTGCACTGATGGAGAGATGTTTC 420  
361 ATCTATGAAAAGCGATGGGTGGGCTGCTTTCATCCCTGCACTGATGGAGAGATGTTTC 420  
421 TGGGCTGCAGCAATTTCTCTGCATTTAAGGGGCCACATCAGCGTATCATTTATGTGAT 480  
421 TGGGCTGCAGCAATTTCTCTGCATTTAAGGGGCCACATCAGCGTATCATTTATGTGAT 480  
481 GTGAACATATCGGTATGTCCTGCACTCATTTGACATCTTTATATCCCTAGTGGGTGG 540  
481 GTGAACATATCGGTATGTCCTGCACTCATTTGACATCTTTATATCCCTAGTGGGTGG 540  
541 CTCTACTCTGTGGCATATCTGATGTGTGTCAGCTATTCGATTTTTTATAGACGTGTGG 600  
541 CTCTACTCTGTGGCATATCTGATGTGTGTCAGCTATTCGATTTTTTATAGACGTGTGG 600  
601 ATCACTGTCCCTTTTGGCCCTGTCAATCCTGTGACATCCGACATGGAGTTTCAAGTGTGG 660  
601 ATCACTGTCCCTTTTGGCCCTGTCAATCCTGTGACATCCGACATGGAGTTTCAAGTGTGG 660  
661 CATGCTAAATACCAAGTCCCTGGCTGGAGACCATTTGAATCAGTTGAATCTTACACTGG 720  
661 CATGCTAAATACCAAGTCCCTGGCTGGAGACCATTTGAATCAGTTGAATCTTACACTGG 720  
721 CTGTAATATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTATCCAGAG 780  
721 CTGTAATATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTATCCAGAG 780

781 GTCTCTCTTATATCTTACGCCAAGCTATGCTGAGGTAAGTCTGCTCTTCTGGCAGCTTTGGG 840  
781 GTCTCTCTTATATCTTACGCCAAGCTATGCTGAGGTAAGTCTGCTCTTCTGGCAGCTTTGGG 840  
841 TGGCCTGGATGAGGCTCTACCCGCCATANGATAGGAGCTATTTGGAGCTTCCACAGACTGG 900  
841 TGGCCTGGATGAGGCTCTACCCGCCATANGATAGGAGCTATTTGGAGCTTCCACAGACTGG 900  
901 AACCAAGTGCCTACGAGGTATCCAGATCCAAAGTAAAGAGAGACAGACATGATCTC 960  
901 AACCAAGTGCCTACGAGGTATCCAGATCCAAAGTAAAGAGAGACAGACATGATCTC 960  
961 CCGATTCGCTTGCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
961 CCGATTCGCTTGCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
1021 TCAGCTGCTGCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 TCAGCTGCTGCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1081 CGGAATATCTACCAAGCTTCTCTTCCAGCAAAATGATCAGACAAAGAAATGTGTGGGTC 1140  
1081 CGGAATATCTACCAAGCTTCTCTTCCAGCAAAATGATCAGACAAAGAAATGTGTGGGTC 1140  
1141 ATGAGGATCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
1141 ATGAGGATCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
1201 ACTGTGTATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
1201 ACTGTGTATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
1261 CTGCTCTGTGTAATCTTCAATCAAGAAACAACTTATGAGGAGCTGCTGCTGCTGCTGCTGCT 1320  
1261 CTGCTCTGTGTAATCTTCAATCAAGAAACAACTTATGAGGAGCTGCTGCTGCTGCTGCTGCT 1320  
1321 TTTGGACATATTCCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1321 TTTGGACATATTCCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1381 TTTCTACCTGGTATTTATCTGACAAAGATGTATATCAATCAGAGGTTCCCATTTTAA 1440  
1381 TTTCTACCTGGTATTTATCTGACAAAGATGTATATCAATCAGAGGTTCCCATTTTAA 1440  
1441 ACTCTCTCATGATGATCTTCAATCTTTTACCAAACTTTGTTTCTTATCTGAGCCAAATAT 1500  
1441 ACTCTCTCATGATGATCTTCAATCTTTTACCAAACTTTGTTTCTTATCTGAGCCAAATAT 1500  
1501 CTATTTGAAAGTGAACCTTGGCTTCCAAATTAAGTGAATTTGATGATGCTGCTGCTGCTGCTGCT 1560  
1501 CTATTTGAAAGTGAACCTTGGCTTCCAAATTAAGTGAATTTGATGATGCTGCTGCTGCTGCTGCT 1560  
1561 CACAGTGAAGAAATGAGAGACATGAGACATCTTCTAGTCAAGAAATGAAGAAATGAAGTAAAT 1620  
1561 CACAGTGAAGAAATGAGAGACATGAGACATCTTCTAGTCAAGAAATGAAGAAATGAAGTAAAT 1620  
1621 GAACCTTGACCTGTGAAACCTTGGCAGAGCTTAACTTCACTTCACTTCACTTCACTTCACTTCA 1680  
1621 GAACCTTGACCTGTGAAACCTTGGCAGAGCTTAACTTCACTTCACTTCACTTCACTTCACTTCA 1680  
1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1741 TGA 1743  
1741 TGA 1743

RESULT 7  
BD012718 1743 bp DNA linear PAT 02-AUG-2002  
LOCUS BD012718  
DEFINITION High-affinity choline transporter.

ACCESSION BD012718  
 VERSION BD012718.1 GI:22092907  
 KEYWORDS WO 0116315-A/2.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciognathi; Muridea; Muridae; Murinae; Rattus.  
 1 (bases 1 to 1743)  
 Hagi, T. and Okuda, T.  
 High-affinity choline transporter  
 Patent: WO 0116315-A 2 08-MAR-2001;  
 JOURNAL JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA  
 OS Rattus norvegicus (rat)  
 COMMENT WO 0116315-A/2  
 PD 08-MAR-2001  
 PF 18-AUG-2000 WO 2000JP005545  
 PR 22-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI  
 TATSUYA HAGA, TAKASHI OKUDA  
 PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC  
 A61K38/17  
 PC A61K45/00, A61P25/28, G01N33/53, A01K67/027  
 CC  
 FH Key Location/Qualifiers  
 FT CDS Location/Qualifiers  
 1..1743  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 ORIGIN  
 Query Match 92.7%; Score 1616.6; DB 2; Length 1743;  
 Best Local Similarity 95.5%; Pred. No. 0;  
 Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Db  
 541 CTCTACTGTGGCATATACGATGTGTAACGATATCTGATTTTATAGAGTGTGG 600  
 601 ATCACTGTCCTTTTGGCCCTGACATCCTGAGTACCGGATGGGATTCACAGCTGG 660  
 602 ATCACTGTCCTTTTGGCCCTGACATCCTGAGTACCGGATGGGATTCACAGCTGG 660  
 603 ATCACTGTCCTTTTGGCCCTGACATCCTGAGTACCGGATGGGATTCACAGCTGG 660  
 661 CATGCTAAATACAGAGTCCCTGGCTGGGAACATTGAATCAGTTGAAGTCAACCTGG 720  
 662 CATGCTAAATACAGAGTCCCTGGCTGGGAACATTGAATCAGTTGAAGTCAACCTGG 720  
 721 CTGTAAATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTATCCAGAG 780  
 722 CTGTAAATTTTCTGTATTTGATGCTGGGTGAATCCATGGCAAGCTTATCCAGAG 780  
 781 GTCTCTCTTCATCCTGACGCACTATGCTCAGTATGCTCTCTCTGACCTTTGGG 840  
 782 GTCTCTCTTCATCCTGACGCACTATGCTCAGTATGCTCTCTCTGACCTTTGGG 840  
 783 GTCTCTCTTCATCCTGACGCACTATGCTCAGTATGCTCTCTCTGACCTTTGGG 840  
 841 TSCCTGTGATGAGCTCTACCCGCAATATGAGAGCTATTTGAGCTTCCACAGACTGG 900  
 842 TSCCTGTGATGAGCTCTACCCGCAATATGAGAGCTATTTGAGCTTCCACAGACTGG 900  
 843 TSCCTGTGATGAGCTCTACCCGCAATATGAGAGCTATTTGAGCTTCCACAGACTGG 900  
 844 TSCCTGTGATGAGCTCTACCCGCAATATGAGAGCTATTTGAGCTTCCACAGACTGG 900  
 901 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTTACAGAGGAGGAGCATGATTTCTC 960  
 902 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTTACAGAGGAGGAGCATGATTTCTC 960  
 961 CCGATGCTTTCAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 962 CCGATGCTTTCAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 1021 TCAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 1022 TCAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 1081 CGGAATATCTACAGCTTCTTCTTCAAGCAAAATGATCAGACAGAAATGCTGCTGCT 1140  
 1082 CGGAATATCTACAGCTTCTTCTTCAAGCAAAATGATCAGACAGAAATGCTGCTGCT 1140  
 1141 ATGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 1142 ATGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
 1201 ACTGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1202 ACTGATGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1261 CTGCTCTGTGTACTCTTCAATCAAGGAGCAACCTTATGAGGAGCTGCTGCTGCTGCT 1320  
 1262 CTGCTCTGTGTACTCTTCAATCAAGGAGCAACCTTATGAGGAGCTGCTGCTGCTGCT 1320  
 1321 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1380  
 1322 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1380  
 1323 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1380  
 1381 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1440  
 1382 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1440  
 1383 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1440  
 1384 TTTGAGCTTCTTCTGAGATTAATCTGAGAGAGGCAATCTTATCTTGAAGCTTAAATC 1440  
 1441 ACTCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1442 ACTCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 1501 CTATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
 1502 CTATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560  
 1561 CACAGTGAAG 1620  
 1562 CACAGTGAAG 1620  
 1621 GAACTTGAAG 1680  
 1622 GAACTTGAAG 1680







Db 1441 ACTCTCATGTTACCTCATCTTACCAACATTTGTTTCCATCTAGCAAGTAT 1500  
QY 1501 CTTATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATGCTGTTCGCAAG 1560  
Db 1501 CTTATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATGCTGTTCGCAAG 1560  
QY 1561 CACAGTGAAGAGACATGAGACACATTTAGTACAGAAATGAAATTCAAATTAAT 1620  
Db 1561 CACAGTGAAGAGACATGAGACACATTTAGTACAGAAATGAAATTCAAATTAAT 1620  
QY 1621 GAACCTTGACCTGTGAACCTTGGCAGACCTTAACCTTCACTTCAACATTAAG 1680  
Db 1621 GAACCTTGACCTGTGAACCTTGGCAGACCTTAACCTTCACTTCAACATTAAG 1680  
QY 1681 GAGGCTCTCTGTTGATGATTTCCAGTCCGAGAGGCTCTGGAGTGAAGTAAATTAACA 1740  
Db 1681 GAGGCTCTCTGTTGATGATTTCCAGTCCGAGAGGCTCTGGAGTGAAGTAAATTAACA 1740  
QY 1741 TGA 1743  
Db 1741 TGA 1743  
RESULT 9  
AB030947 4904 bp mRNA linear ROD 03-FEB-2000  
LOCUS AB030947  
DEFINITION Rattus norvegicus mRNA for high-affinity choline transporter CHTL,  
complete cds.  
ACCESSION AB030947  
VERSION AB030947.1 GI:6863033  
KEYWORDS choline transporter; high-affinity choline transporter CHTL.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Rattus.  
1 (sites)  
REFERENCE Okuda, T., Haga, T., Kanai, Y., Endou, H., Iehihara, T. and Katsura, I.  
AUTHORS Identification and characterization of the high-affinity choline  
TITLE transporter  
JOURNAL Nat. Neurosci. 3 (2), 120-125 (2000)  
PUBMED 10649566  
REFERENCE 2 (bases 1 to 4904)  
AUTHORS Okuda, T.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-1999) Takashi Okuda, University of Tokyo, Faculty  
of Medicine, Department of Neurochemistry, Hongo 7-3-1, Bunkyo-ku  
113-0033, Japan (E-mail: okuda@n.u-tokyo.ac.jp, Tel: +81-3-5841-3560,  
Fax: +81-3-3814-8154)  
COMMENT Sequence updated (11-Jan-2000).  
FEATURES  
Source location/Qualifiers  
1..4904  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/clone="CHTL"  
/tissue\_type="spinal cord"  
/clone\_lib="rat spinal cord cDNA library"  
/dev\_stage="adult"  
224..1966  
/codon\_start=1  
/product="high-affinity choline transporter CHTL"  
/protein\_id="BAA90484.1"  
/translation="MEFHEGLVAIIIFYLLIFLVGIWAMKTKNSGAERSEALIV  
GGRDIGLVGFTMTATWGGYINGTAAVYGGGLMAQAIFGYSLSLIGLFF  
APRMSKGVMTLPQOIIYGRMGSLFI PALMGEMFAAIFSLGATISVILIDV  
VNIYSVSLAILIYLVGLYSVAITDVOLFCPIGAMISVPLSHPAVDTIGFT  
AVHAKYSPMLCTTISVEYVTLNDNLLMGISIPQAVFORVLSSSATYACOVSLF  
AAPGCIYVALPAICIGATGASIDMNQTAIGFPDKTKREADNLIPIVLOYLCPVYSF  
FELGAVSAVMSADSSILSASSMPARNIYOUSFRONASDKETIWMVRITVFAGSA

ORIGIN  
Query Match 92.7%; Score 1616.6; DB 6; Length 4904;  
Best Local Similarity 95.5%; Pred. No. 0;  
Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
TAMALLTKTVGLWYLSLDVYIIIFPOLCVLFIKGNITGAVAGYIFGLFRTGG  
EPYALQPLIFPYGYPDKNGIYNRFPEKILSMSTFETNLCVSLAYLSEGLP  
PDLIDFADVSHSEENMDKTLIVNENNIKLNELAPVKRQSLTISSTFTNKALLDV  
DSSPESGIEDNLIQ"  
QY 1 ATGCCCTTTCATGTGAGAGAGCTGAGTATATTCCTCTTACTCTTATATTTCTG 60  
Db 224 ATGCCCTTTCATGTGAGAGAGCTGAGTATATTCCTCTTACTCTTATATTTCTG 283  
QY 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGGCGCAACCCAGAAAGGCGATGAA 120  
Db 284 GTTGAATATGAGGCTGATGAGAAACCAAAACAGGCGTAAATGAGAAAGCGAGCAA 343  
QY 121 GCCATCATATGTCGGGGCCGATGACATTTGTTGTTGTTGTTTACATGACAGCC 180  
Db 344 GCCATCATATGTTGGGGCCGATGACATTTGTTGTTGTTGTTTACATGACAGCC 403  
QY 181 ACCGTGGTTGAGAGAGCTACATCAATGGGACAGAAAGCGTGTATGGCCAGTTGT 240  
Db 404 ACCGTGGTTGAGAGAGCTACATCAATGGGACAGCTGAAGAGCTTATGAGCCAGTTGT 463  
QY 241 GGCTTAGCTTGGGCTCATGACCCCATTTGATATTCCTGAGTCTAATTTAGGTGCTG 300  
Db 464 GGCTTAGCTTGGGCTCATGACCCCATTTGATATTCCTGAGTCTAATTTAGGTGCTG 523  
QY 301 TTTTTCGCAAACTATGCGTTCGCAAGGGATATGACTATGTAGACCATTCAAACAG 360  
Db 524 TTTTTCGCAAACTATGCGTTCGCAAGGGATATGACTATGTAGACCATTCAAACAG 583  
QY 361 ATCTATGAAAGCGATGGGTGGGCTCTTCATCTCTGCACTGATGGAGAGATGTTTC 420  
Db 584 ATCTATGAAAGCGATGGGTGGGCTCTTCATCTCTGCACTGATGGAGAGATGTTTC 643  
QY 421 TGGGCGTGAAGAAATTTCTCTGCTATGAGGGGCAACATGAGGCTGATTCATGATGAT 480  
Db 644 TGGGCGTGAAGAAATTTCTCTGCTATGAGGGGCAACATGAGGCTGATTCATGATGAT 703  
QY 481 GTGAACATATCGGATATGCTCTGCACTATGCTCATTTATACCTAGTGGTGG 540  
Db 704 GTGAACATATCGGATATGCTCTGCACTATGCTCATTTATACCTAGTGGTGG 763  
QY 541 CTCTACTCTGTCATATATCTGATGTTGTCAGCTATTTCTGCATTTTATAGACTGTG 600  
Db 764 CTCTACTCTGTCATATATCTGATGTTGTCAGCTATTTCTGCATTTTATAGACTGTG 823  
QY 601 ATCAGTGCCTTTTGGCCCTGTCACATCTGTCAGTACCGACATCGGATTCACAGCTGTG 660  
Db 824 ATCAGTGCCTTTTGGCCCTGTCACATCTGTCAGTACCGACATCGGATTCACAGCTGTG 883  
QY 661 CATGCTAAATACAGAGTCCCTGGCTGGAGAACATTAATCACTTGAAGCTTACCTGG 720  
Db 884 CATGCTAAATACAGAGTCCCTGGCTGGAGAACATTAATCACTTGAAGCTTACCTGG 943  
QY 721 CTTGATATTTCTGTATATGATGCTGGGTGAATCCCATGCGACCTTATCTCCAGAG 780  
Db 944 CTTGATATTTCTGTATATGATGCTGGGTGAATCCCATGCGACCTTATCTCCAGAG 1003  
QY 781 GTCTCTCTTCAATCCGACGACCTATGCTCAGGTATCTCTTCTGCGAGCTTTGGG 840  
Db 1004 GTCTCTCTTCAATCCGACGACCTATGCTCAGGTATCTCTTCTGCGAGCTTTGGG 1063  
QY 841 TGCCTGTGATGCTCTTACCGCCCATATGATGATGAGACTATTTGAGACTTCCACAGCTGG 900  
Db 1064 TGCCTGTGATGCTCTTACCGCCCATATGATGATGAGACTATTTGAGACTTCCACAGCTGG 1123  
QY 901 AACCGAGCTGCTAGGGGATCCAGATCCCAAGACTAAGAGAGAGCAGACATGATTCTC 960  
Db 1124 AACCGAGCTGCTAGGGGATCCAGATCCCAAGACTAAGAGAGAGCAGACATGATTCTC 1183

QY 961 CGGATCGTCTGACGACCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1020  
DB 1184 CCGATGTTCTTACAGTACCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1243  
QY 1021 TCAGCTGCTGATGCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1080  
DB 1244 TCTGCTGCTGATGCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1303  
QY 1081 CGGATATCTACGACCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1140  
DB 1304 CCGATATCTACGACCTCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1363  
QY 1141 ATGAGGATCATCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1200  
DB 1364 ATGAGGATCATCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1423  
QY 1201 ACTGCTGATGAGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1260  
DB 1424 ACTGCTGATGAGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1483  
QY 1261 CTGCTCTGCTGATCTCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1320  
DB 1484 CTGCTCTGCTGATCTCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1543  
QY 1321 TTTGGACATCTCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1380  
DB 1544 TTTGGACATCTCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1603  
QY 1381 TTTGATCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1440  
DB 1604 TTTGATCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1663  
QY 1441 ACTCTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1500  
DB 1664 ACTCTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1723  
QY 1501 CTATTTGAAAGTGAACCTCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1560  
DB 1724 CTATTTGAAAGTGAACCTCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1783  
QY 1561 CACAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
DB 1784 CACAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1843  
QY 1621 GAACTTGACCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1680  
DB 1844 GAACTTGACCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1903  
QY 1681 GAGGCTCTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1740  
DB 1904 GAGGCTCTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 1963  
QY 1741 TGA 1743  
DB 1964 TGA 1966

RESULT 10  
AR268949 1743 bp DNA linear PAT 10-APR-2003  
LOCUS AR268949  
DEFINITION Sequence 1 from patent US 6500643.  
ACCESSION AR268949  
VERSION AR268949.1 GI:29699686  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

FEATURES  
source  
Location/Qualifiers  
1..1743  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 79.1%; Score 1378.2; DB 2; Length 1743;  
Best Local Similarity 86.9%; Pred.No. 0; Mismatches 228; Indels 0; Gaps 0;  
Matches 1515; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

QY 1 ATGCTTTCCATGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 60  
DB 1 ATGCTTTCCATGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 60  
QY 61 GTTGAATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 120  
DB 61 GTTGAATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 120  
QY 121 GGCATCATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 180  
DB 121 GGCATCATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 180  
QY 181 ACTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 240  
DB 181 ACTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 240  
QY 241 GGTCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 300  
DB 241 GGTCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 300  
QY 301 TTTTGGGAAACCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 360  
DB 301 TTTTGGGAAACCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 360  
QY 361 ATCTATGAAACCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 420  
DB 361 ATCTATGAAACCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 420  
QY 421 TGGGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 480  
DB 421 TGGGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 480  
QY 481 GGAACATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 540  
DB 481 GGAACATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 540  
QY 541 CTCTACTCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 600  
DB 541 CTCTACTCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 600  
QY 601 ATCACTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 660  
DB 601 ATCACTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 660  
QY 661 CATGCTAATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 720  
DB 661 CATGCTAATATGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 720  
QY 721 CTGATTAATTTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 780  
DB 721 CTGATTAATTTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 780  
QY 781 GTCTCTCTTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 840  
DB 781 GTCTCTCTTCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 840  
QY 841 TGCTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 900  
DB 841 TGCTGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 900  
QY 901 AACGAGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 960  
DB 901 AACGAGCTGCTGCTGCTGCTGCTGCTGATCATCTCTCTTCTGAGCTTGCTGT 960

```

QY 961 CCGATGCTTCTGAGTACCTGCTCCCTGATACATCTCTTTGGGGCTGGTCTGTT 1020
DB 961 CCAATGTTCTGAGATCTCTGCTGCTGATATTTCTTTTGGCTTGGTGCAGTT 1020
QY 1021 TCAGTGTCTGATGCTCTCACTGAGTCACTGTCATCTGTGGCGATTTATGTTGCT 1080
DB 1021 TCAGTGTCTGATGCTCTCACTGAGTCACTGTCATCTGTGGCGATTTATGTTGCT 1080
QY 1081 CCGAATATCTACAGCTTCTCTTCAAGCAAAATGATCAGACAGAGAAATGTGTGGTTC 1140
DB 1081 CCGAATATCTACAGCTTCTCTTCAAGCAAAATGATCAGACAGAGAAATGTGTGGTTC 1140
QY 1141 ATGAGATCACTGTGCTTGTGTTGAGAGATCTGCAACAGCCATGCTTGTGAGAGAG 1200
DB 1141 ATGAGATCACTGTGCTTGTGTTGAGAGATCTGCAACAGCCATGCTTGTGAGAGAG 1200
QY 1201 ACTGTGTATGGGCTCTGATACCTGAGCTTGTCTACATCATCTTCCGACAG 1260
DB 1201 ACTGTGTATGGGCTCTGATACCTGAGCTTGTCTACATCATCTTCCGACAG 1260
QY 1261 CTGCTCTGTGTACTCTTCTCAACAGAACACATTTAGGGCACTTGTGTTATTT 1320
DB 1261 CTGCTCTGTGTACTCTTCTCAACAGAACACATTTAGGGCACTTGTGTTATTT 1320
QY 1321 TTGGAATATCTGAGAAATTAATCTGAGAGAGAGCCATCTATCTTGAAGCCCTTAATC 1380
DB 1321 TTGGAATATCTGAGAAATTAATCTGAGAGAGAGCCATCTATCTTGAAGCCCTTAATC 1380
QY 1381 TTCTACCTGCTTATTAATCTGACAAAGATGTATATCAATCAGAGTTCCCATTTAAA 1440
DB 1381 TTCTACCTGCTTATTAATCTGACAAAGATGTATATCAATCAGAGTTCCCATTTAAA 1440
QY 1441 ACTCTCTCATGATTAATCTTCAATCTTCTTACCAATTTGTTCTTATCTAGCAAGTAT 1500
DB 1441 ACTCTCTCATGATTAATCTTCAATCTTCTTACCAATTTGTTCTTATCTAGCAAGTAT 1500
QY 1501 CTATTGGAAGTGAAGCTTGGCTCCAAATTAATGATGATTTGATCTGTGTCAGAG 1560
DB 1501 CTATTGGAAGTGAAGCTTGGCTCCAAATTAATGATGATTTGATCTGTGTCAGAG 1560
QY 1561 CACAGTGAAGAGAACATGACAAACATCTTATGTCAGAAATGAAATATCAAAATTAAT 1620
DB 1561 CACAGTGAAGAGAACATGACAAACATCTTATGTCAGAAATGAAATATCAAAATTAAT 1620
QY 1621 GAATCTGCACTGTGAAACCTGGGCAAGCTTAACCTCACTTCACTTACCAATTAAG 1680
DB 1621 GAATCTGCACTGTGAAACCTGGGCAAGCTTAACCTCACTTCACTTACCAATTAAG 1680
QY 1681 GAGGCGCTCTGATGATTAATCTGAGTCCGAGAGGGGCTGGGAGCTGAAGTAATTTACA 1740
DB 1681 GAGGCGCTCTGATGATTAATCTGAGTCCGAGAGGGGCTGGGAGCTGAAGTAATTTACA 1740
QY 1741 TGA 1743
DB 1741 TGA 1743

```

```

RESULT 11
BD012719 1743 bp DNA linear PAT 02-AUG-2002
LOCUS High-affinity choline transporter.
DEFINITION BD012719
ACCESSION BD012719.1 GI:22092908
VERSION WO 0116315-A/3.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1743)
Haga,T. and Okuda,T.
REFERENCE
AUTHORS Haga,T. and Okuda,T.
TITLE High-affinity choline transporter

```

```

JOURNAL Patent: WO 0116315-A 3 08-MAR-2001;
JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT OS Homo sapiens (human)
PN WO 0116315-A/3
PD 08-MAR-2001
PF 18-AUG-2000 WO 2000JP005545
PR 27-AUG-1999 JP 99P 240642,27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
A61K38/17,
PC A61K45/00, A61P25/28, G01N33/53, A01K67/027
CC FH Key Location/Qualifiers
FT CDS (1)..(1743).
FEATURES
source 1..1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 78.9%; Score 1375; DB 2; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
QY 1 ATGACCTTTCATGTGGAAGAGCTGTAGCTATATATCTCTTCACTCTTATATTTCTG 60
DB 1 ATGACCTTTCATGTGGAAGAGCTGTAGCTATATATCTCTTCACTCTTATATTTCTG 60
QY 61 GTTGAATATGAGGCTGTCATGAGAAACCAAAACAGGGGCAACCCAGAGCGCAGTGA 120
DB 61 GTTGAATATGAGGCTGTCATGAGAAACCAAAACAGGGGCAACCCAGAGCGCAGTGA 120
QY 121 GCCATATATGTGGGGGCGCTGACATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 121 GCCATATATGTGGGGGCGCTGACATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 121 GCCATATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 121 GCCATATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 ACCGTGGGTGGAAGAGGCTGATCATCAATGGGAGACAGAGAGCTGATGGGCGAGTGT 240
DB 181 ACCGTGGGTGGAAGAGGCTGATCATCAATGGGAGAGAGAGCTGATGGGCGAGTGT 240
QY 181 ACCGTGGGTGGAAGAGGCTGATCATCAATGGGAGAGAGAGCTGATGGGCGAGTGT 240
DB 181 ACCGTGGGTGGAAGAGGCTGATCATCAATGGGAGAGAGAGCTGATGGGCGAGTGT 240
QY 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAAATTTAGGTGCTG 300
DB 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAAATTTAGGTGCTG 300
QY 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAAATTTAGGTGCTG 300
DB 241 GGTCTAGCTTGGGCTCATGACCCATTTGATATTTCTGAGTCTAAATTTAGGTGCTG 300
QY 301 TTTTTCGCAAACTTATGCTTCCAAAGGATATGTACTATGTATAGACCATTTCAACAG 360
DB 301 TTTTTCGCAAACTTATGCTTCCAAAGGATATGTACTATGTATAGACCATTTCAACAG 360
QY 301 TTTTTCGCAAACTTATGCTTCCAAAGGATATGTACTATGTATAGACCATTTCAACAG 360
DB 301 TTTTTCGCAAACTTATGCTTCCAAAGGATATGTACTATGTATAGACCATTTCAACAG 360
QY 361 ATCTATGGAAGAGAGAGGAGGAGCTCTCTTATTTCTGACATGATGGAGAAATGTTTC 420
DB 361 ATCTATGGAAGAGAGAGGAGGAGCTCTCTTATTTCTGACATGATGGAGAAATGTTTC 420
QY 421 TGGGCTGCAAGCAATTTCTCTGCTTTGGAGGACACATCAGGGTATCATTCATGAT 480
DB 421 TGGGCTGCAAGCAATTTCTCTGCTTTGGAGGACACATCAGGGTATCATTCATGAT 480
QY 481 GTGACATATTCGGTCAATTTCTCTGACATCTTGGCATTTCTTATACCTAGTGGGTGG 540
DB 481 GTGACATATTCGGTCAATTTCTCTGACATCTTGGCATTTCTTATACCTAGTGGGTGG 540
QY 541 CTCTACTGTGGGCAATATGATGATGTCAGAGTATTCGATTTTATAGAGTCTG 600
DB 541 CTCTACTGTGGGCAATATGATGATGTCAGAGTATTCGATTTTATAGAGTCTG 600
QY 541 CTCTACTGTGGGCAATATGATGATGTCAGAGTATTCGATTTTATAGAGTCTG 600
DB 541 CTCTACTGTGGGCAATATGATGATGTCAGAGTATTCGATTTTATAGAGTCTG 600
QY 601 ATGAGTGTCTTTTGGCTGCAATCTGCAAGTCAACGAGATGATTCACAGCTGTG 660
DB 601 ATGAGTGTCTTTTGGCTGCAATCTGCAAGTCAACGAGATGATTCACAGCTGTG 660
QY 601 ATGAGTGTCTTTTGGCTGCAATCTGCAAGTCAACGAGATGATTCACAGCTGTG 660
DB 601 ATGAGTGTCTTTTGGCTGCAATCTGCAAGTCAACGAGATGATTCACAGCTGTG 660
QY 661 CATGCTAAATATCAGAGTCCCTGGGAGGAGACATTTGATGATGATGATGATGATGAT 720
DB 661 CATGCTAAATATCAGAGTCCCTGGGAGGAGACATTTGATGATGATGATGATGATGAT 720

```

QY 721 CTTGATATTTTCTGTATTTGATGCTGGGTAATCCATGGCAAGCTTCCAGAG 780  
DB 721 CTTGATATTTTCTGTGTGTGATGCTGGGTAATCCATGGCAAGCTTCCAGAG 780  
QY 781 GTCCTCTCTATCTCCAGACCTATGCTGAGTACGTCTCTCTGACCTTTGGG 840  
DB 781 GTCCTCTCTCTCTCTCCAGACCTATGCTGAGTACGTCTCTCTGACCTTTGGG 840  
QY 841 TGCCTGTGATGGCTCTACCCGCCATATGATGAGTATGAGCTTTCACAGACTGG 900  
DB 841 TGCCTGTGATGGCTCTACCCGCCATATGAGTATGAGCTTTCACAGACTGG 900  
QY 901 AACGAGCTGCTAGGGGATCCAGATCCCAAGATTAAGAGAGAGACATGATCTC 960  
DB 901 AACGAGCTGCTAGGGGATCCCAATCCCAAGATTAAGAGAGAGACATGATCTC 960  
QY 961 CCGATGCTTCTGACGACTCTGACCTGCTGATCTCTCTTGGGCTTGGTCTGT 1020  
DB 961 CCGATGCTTCTGACGACTCTGACCTGCTGATCTCTCTTGGGCTTGGTCTGT 1020  
QY 1021 TCAGCTGCTGATGCTGCTGACCTGCTGATCTCTGCTGCTGCTGCTGCT 1080  
DB 1021 TCAGCTGCTGATGCTGCTGACCTGCTGATCTCTGCTGCTGCTGCTGCT 1080  
QY 1081 CGGAATATCTACGACTTCTCTGACCAAAATGCAATGCAAGAAATTTGTGGT 1140  
DB 1081 CGGAATATCTACGACTTCTCTGACCAAAATGCTTCCGCAAAAGAAATTTGGGTT 1140  
QY 1141 ATGAGATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1141 ATGAGATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
QY 1261 CTGCTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
DB 1261 CTGCTGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320  
QY 1321 TTTGACTATTTCTGAGATTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
DB 1321 TTTGACTATTTCTGAGATTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
QY 1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
DB 1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
QY 1441 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 1441 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
QY 1501 CTTATTTGAAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATGCTGCT 1560  
DB 1501 CTTATTTGAAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATGCTGCT 1560  
QY 1561 CACAGTGAAG 1620  
DB 1561 CACAGTGAAG 1620  
QY 1621 GAACTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
DB 1621 GAACTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY 1681 GAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
DB 1681 GAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
QY 1741 TGA 1743  
DB 1741 TGA 1743

RESULT 12  
CQ722844 1743 bp DNA linear PAT 03-FEB-2004  
LOCUS CQ722844  
DEFINITION Sequence 8778 from Patent WO02068579.  
ACCESSION CQ722844  
VERSION CQ722844.1 GI:42283701  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 8778 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
Location/Qualifiers  
1..1743  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 78.9%; Score 1375; DB 2; Length 1743;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;  
QY 1 ATGCTTTCCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
DB 1 ATGCTTTCCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
QY 61 GTTGAATATGAGGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
DB 61 GTTGAATATGAGGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 121 GGCATCATATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
DB 121 GGCATCATATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 181 ACCTGGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
DB 181 ACCTGGGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
QY 241 GGTCTAGCTTGGGCTCATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
DB 241 GGTCTAGCTTGGGCTCATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
QY 301 TTTTTCGAGAAACCTATGCTTCCAGAGGATATGATGATGATGATGATGAT 360  
DB 301 TTTTTCGAGAAACCTATGCTTCCAGAGGATATGATGATGATGATGATGAT 360  
QY 361 ATCTATGAGAAAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
DB 361 ATCTATGAGAAAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
QY 421 TGGGCTGAGCAATTTCTCTGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
DB 421 TGGGCTGAGCAATTTCTCTGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 481 GTGAACATATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 540  
DB 481 GTGAACATATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 541 CTCTACTGCTGAGGAGATATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 541 CTCTACTGCTGAGGAGATATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 601 ATCAGTGTCCCTTTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 601 ATCAGTGTCCCTTTTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 660

```

OY 661 CATGTAATATACAGAGTCCCTGGCTGGGAACCATTAAGTGAAGTCAACCTGG 720
DB 661 CATGCAATATACCAAAAGCCGCTGGGACCTGTTGACTCATCTGAAGTCTACTCTTGG 720
OY 721 CTGTAATATTTTCTGTATATGATGCTGGGTGGAATCCCATGGCAAGCCCTACTTCCAGAG 780
DB 721 CTGTAATATTTTCTGTATATGATGCTGGGTGGAATCCCATGGCAAGCCCTACTTCCAGAG 780
OY 781 GTTCCTCTCTTCACTCTGAGCCCACTATGCTGAGTACTGCTCTTCTGGCAGCTTTGGG 840
DB 781 GTTCCTCTCTTCTCTGAGCCCACTATGCTGAGTACTGCTCTTCTGGCAGCTTTGGG 840
OY 841 TGCCGTGATGAGTCTCTACCCGCCCATATGATAGAGTATGAGTCTCCAGAGCTGG 900
DB 841 TGCCGTGATGAGTCTCTACCCGCCCATATGATAGAGTATGAGTCTCCAGAGCTGG 900
OY 901 AACGAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAAGAGAGAGACATGATCTC 960
DB 901 AACGAGACTGATATGGGCTTCAGATCCCAAGACTTAAAGAGAGAGACATGATCTC 960
OY 961 CGGATGCTTCTGAGTACTCTGCTGCTGTGTCATCTCTTCTTGGGCTTGGTCTGT 1020
DB 961 CGGATGCTTCTGAGTACTCTGCTGCTGTGTCATCTCTTCTTGGGCTTGGTCTGT 1020
OY 1021 TCAGTGTGCTGATGCTCTCAGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TCAGTGTGCTGATGCTCTCAGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 1080
OY 1081 CGGAATATCTTACAGCTTCTTCAAGCAAAATGATCAGACAGAGAAATGCTGGTCTC 1140
DB 1081 CGGAATATCTTACAGCTTCTTCAAGCAAAATGATCAGACAGAGAAATGCTGGTCTC 1140
OY 1141 ATGAGATATCAGTGTGCTGTGCTGAGAGATCTGCAAGAGCCATGCTTCTGCAAGAG 1200
DB 1141 ATGAGATATCAGTGTGCTGTGCTGAGAGATCTGCAAGAGCCATGCTTCTGCAAGAG 1200
OY 1201 ACTGTGTATGAGGCTCGGATCTGAGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1260
DB 1201 ACTGTGTATGAGGCTCGGATCTGAGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1260
OY 1261 CTGCTCTGTGATCTCTTCAACAAAGAAACCAACATTAATGAGGAGCTGCTGATATAT 1320
DB 1261 CTGCTCTGTGATCTCTTCAACAAAGAAACCAACATTAATGAGGAGCTGCTGATATAT 1320
OY 1321 TTTGAGATTAATCTGAGAAATTAAGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1380
DB 1321 TTTGAGATTAATCTGAGAAATTAAGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1380
OY 1381 TTTGAGATTAATCTGAGAAATTAAGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1440
DB 1381 TTTGAGATTAATCTGAGAAATTAAGAGAGAGCCATATCTATCTTGAAGCCCTTAATC 1440
OY 1441 ACTCTCTCATGATTAATCTGAGAAATTAAGAGAGCCATATCTATCTTGAAGCCCTTAAT 1500
DB 1441 ACTCTCTCATGATTAATCTGAGAAATTAAGAGAGCCATATCTATCTTGAAGCCCTTAAT 1500
OY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATTAATGATGCTGTGCTGCAAG 1560
DB 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTAAGATTAATGATGCTGTGCTGCAAG 1560
OY 1561 CACAGTGAAGAGAAATGAGCAAGACCATTAAGTGAAGATTAAGAAATTAAGTGAAGT 1620
DB 1561 CACAGTGAAGAGAAATGAGCAAGACCATTAAGTGAAGATTAAGAAATTAAGTGAAGT 1620
OY 1621 GAATTTGCACTGTGAAACCTTGGCAAGCTTAACCTTCAAGTTCAACCTTAAG 1680
DB 1621 GAATTTGCACTGTGAAACCTTGGCAAGCTTAACCTTCAAGTTCAACCTTAAG 1680
OY 1681 GAGGCTCTCTGATGATTAATGATGCTGAGAGAGGCTTGGGAGCTGAGAGATTAATTA 1740
DB 1681 GAGGCTCTCTGATGATTAATGATGCTGAGAGAGGCTTGGGAGCTGAGAGATTAATTA 1740

```

```

OY 1741 TGA 1743
DB 1741 TGA 1743

RESULT 13
E49871 1743 bp DNA linear PAT 27-AUG-2002
LOCUS High-affinity choline transporter.
ACCESSION E49871.1 GI:22554902
VERSION E49871.1
KEYWORDS JP 2001136976-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1743)
AUTHORS Haga, T. and Okuda, T.
TITLE High-affinity choline transporter
JOURNAL Patent: JP 2001136976-A 3 22-MAY-2001;
SCIENCE & TECH AGENCY
COMMENT OS Homo sapiens (human)
PN JP 2001136976-A/3
PD 22-MAY-2001
PE 27-DEC-1999 JP 1999368991
PI TATSUYA HAGA, TAKASHI OKUDA
PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,
PC C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1743
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.9%; Score 1375; DB 2; Length 1743;
Best Local Similarity 86.8%; Pred. No. 0;
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

OY 1 ATGCTTTCCATGAGTGAAGAGACTGGAGTATATCTCTTACTCTCTTAATTTCTG 60
DB 1 ATGCTTTCCATGAGTGAAGAGACTGGAGTATATCTCTTACTCTCTTAATTTCTG 60
OY 61 GTTGAATATAGGCTGCATGAGAAACCAAAACAGCGCAACCCAGAGAGCGCAGTGA 120
DB 61 GTTGAATATAGGCTGCATGAGAAACCAAAACAGCGCAACCCAGAGAGCGCAGTGA 120
OY 121 GCCATCATATGTCGGGGGCGGTGACATTTGGTTGTTGGTTGTTTACATGACAGCC 180
DB 121 GCCATCATATGTCGGGGGCGGTGACATTTGGTTGTTGGTTGTTTACATGACAGCC 180
OY 181 ACCTGGGTGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGT 240
DB 181 ACCTGGGTGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGT 240
OY 241 GGCTAGCTTGGGCTCATGACCACTTGGATATCTGAGTCTAATTTAGTGGTCTG 300
DB 241 GGCTAGCTTGGGCTCATGACCACTTGGATATCTGAGTCTAATTTAGTGGTCTG 300
OY 301 TTTTGGGAAACCTATGCTTCCAAAGGATATGACTATGATTAAGCCATTCAAACG 360
DB 301 TTTTGGGAAACCTATGCTTCCAAAGGATATGACTATGATTAAGCCATTCAAACG 360
OY 361 ATCTATGAAAGCGATGGGCTGCTCTTCACTGCACTGATGGAGAGATGTTTC 420
DB 361 ATCTATGAAAGCGATGGGCTGCTCTTCACTGCACTGATGGAGAGATGTTTC 420
OY 421 TGGGCTGAGCAATTTTCTCTGATTAAGGGCCACATCAGGTGATCATTTGATGAT 480
DB 421 TGGGCTGAGCAATTTTCTCTGATTAAGGGCCACATCAGGTGATCATTTGATGAT 480

```

Db 421 TGGGTCGACGAATTTCTCTCTTGGGAGCCACATCAGCTATCATGATGAT 480  
 Qy 481 GTGAACATATCGGTATGTCTCTGACACTCATTTGCAATCTTTATATACCTTATGTTGGTGG 540  
 Db 481 ATGCACATTTCTGTCATCTCTGACACTCATTTGCACTCTGTACACACTGGTGGAGGG 540  
 Qy 541 CTCTACTCTGGGCAATATACGATGATGTTGTCGACGATTTCTGATTTTATAGACCTGG 600  
 Db 541 CTCTACTCTGGGCAATATACGATGATGTTGTCGACGATTTCTGATTTTATAGACCTGG 600  
 Qy 601 ATCAATGTCCTTTGGCCCTGTCATCTCTGACGATCACGACATCGAATTCACAGCTGTG 660  
 Db 601 ATCAATGTCCTTTGGCCCTGTCATCTCTGACGATCACGACATCGAATTCACAGCTGTG 660  
 Qy 661 CATGCTAAATACAGAGTCCCTGCTGGAGAACATTGAATCAGTTGAAGTTACACTGG 720  
 Db 661 CATGCTAAATACAGAGTCCCTGCTGGAGAACATTGAATCAGTTGAAGTTACACTGG 720  
 Qy 721 CTGTGATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCCAGAG 780  
 Db 721 CTGTGATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCCAGAG 780  
 Qy 781 GTCTCTCTGATCTCTGACCCACTATGCTCAGTACTGCTCTCTGCGACGCTTTGGG 840  
 Db 781 GTCTCTCTGATCTCTGACCCACTATGCTCAGTACTGCTCTCTGCGACGCTTTGGG 840  
 Qy 841 TGCTGTGATGCTCTTACCCGCGCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900  
 Db 841 TGCTGTGATGCTCTTACCCGCGCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900  
 Qy 901 AACCGACGCTGATGGGATCCAGATCCCAAGACTAAGAGAGAGACATGATTC 960  
 Db 901 AACCGACGCTGATGGGATCCAGATCCCAAGACTAAGAGAGAGACATGATTC 960  
 Qy 961 CGGATGCTTCTGACGATCTCTGCTGTATCATCTCTCTTGGGCTGTGCTGTT 1020  
 Db 961 CGGATGCTTCTGACGATCTCTGCTGTATCATCTCTCTTGGGCTGTGCTGTT 1020  
 Qy 1021 TCAGTCTGCTGATCTCTGACGATCTCTGCTGTATCATCTCTCTTGGGCTGTGCTGTT 1080  
 Db 1021 TCAGTCTGCTGATCTCTGACGATCTCTGCTGTATCATCTCTCTTGGGCTGTGCTGTT 1080  
 Qy 1081 CGGATATCTACGACCTTCTCTGACGAAATGCAAGAGAGAAATGTTGGGCT 1140  
 Db 1081 CGGATATCTACGACCTTCTCTGACGAAATGCAAGAGAGAAATGTTGGGCT 1140  
 Qy 1141 ATGAGGATCACTGCTGTGTGTTCCGAGCATCTGCAACAGCCATGCTTGTGTCAGAG 1200  
 Db 1141 ATGAGGATCACTGCTGTGTGTTCCGAGCATCTGCAACAGCCATGCTTGTGTCAGAG 1200  
 Qy 1201 ACTGTGTATGGGCTCTGCTGACCTGAGCTGTGACCTTGTCTATCATCATCTTCCACAG 1260  
 Db 1201 ACTGTGTATGGGCTCTGCTGACCTGAGCTGTGACCTTGTCTATCATCATCTTCCACAG 1260  
 Qy 1261 CTGCTGTGTATCTCTTCTGATCAAAAGAACCACTTATGAGGAGCTGCTGTATAT 1320  
 Db 1261 CTGCTGTGTATCTCTTCTGATGAGAACCACTTATGAGGAGCTGCTGTATAT 1320  
 Qy 1321 TTTGGACTATTTCTGAGATTACTGAGAGAGACCATCTATCTTACTTGGAGCCCTTAAATC 1380  
 Db 1321 TTTGGACTATTTCTGAGATTACTGAGAGAGACCATCTATCTTACTTGGAGCCCTTAAATC 1380  
 Qy 1381 TTTCTACCTGCTGTTATTAATCTGACAAAGATGATATACATACAGAGGTTCCCATTTAA 1440  
 Db 1381 TTTCTACCTGCTGTTATTAATCTGACAAAGATGATATACATACAGAGGTTCCCATTTAA 1440  
 Qy 1441 ACTGCTCCATGCTGATACCTGATCTTTTACCAACATTTGTTCTTATATACCAAGAT 1500  
 Db 1441 ACTGCTCCATGCTGATACCTGATCTTTTACCAACATTTGTTCTTATATACCAAGAT 1500  
 Qy 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGATTTGATGCTGTGTCAGAG 1560  
 Db 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGATTTGATGCTGTGTCAGAG 1560

Qy 1561 CACAGTGAAGAACATGACAGAACCAATTTACTAGCAAAATGAAATATCAATTTAAT 1620  
 Db 1561 CACAGTGAAGAACATGACAGAACCAATTTACTAGCAAAATGAAATATCAATTTAAT 1620  
 Qy 1621 GAACCTGACCTGTGAAACCTCGGACAGCCTTAACCCCTCACTTCAACTTCAACAAATAG 1680  
 Db 1621 GAACCTGACCTGTGAAACCTCGGACAGCCTTAACCCCTCACTTCAACTTCAACAAATAG 1680  
 Qy 1681 GAGGCCCTCTGATGTTGATTTCCAGTCCGAGAGGGCTCTGGAGCTGAAGATTAATTTACAA 1740  
 Db 1681 GAGGCCCTCTGATGTTGATTTCCAGTCCAGAGGGCTCTGGAGCTGAAGATTAATTTACAG 1740  
 Qy 1741 TGA 1743  
 Db 1741 TGA 1743  
 RESULT 14  
 AF276871 1743 bp mRNA linear PRI 27-NOV-2000  
 LOCUS AF276871 Homo sapiens high affinity choline transporter (SLC5A7) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF276871  
 VERSION AF276871.1 GI:10998441  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo  
 1 (bases 1 to 1743)  
 Apparsundaram,S., Ferguson,S.M., George,A.L. Jr. and Blakely,R.D.  
 Molecular cloning of a human, hemicholinium-3-sensitive choline  
 transporter  
 JOURNAL Biochem. Res. Commun. 276 (3), 862-867 (2000)  
 PUBMED 11027560  
 REFERENCE 2 (bases 1 to 1743)  
 Apparsundaram,S., Ferguson,S.M. and Blakely,R.D.  
 Direct Submission  
 Submitted (09-JUN-2000) Department of Pharmacology and Center for  
 Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at  
 Pierce, Nashville, TN 37232-6420, USA  
 FEATURES  
 source location/Qualifiers  
 1..1743  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2q12"  
 1..1743  
 /gene="SLC5A7"  
 1..1743  
 /gene="SLC5A7"  
 /note="hCHT: solute carrier family 5 member 7"  
 /codon\_start=1  
 /product="high affinity choline transporter"  
 /protein\_id="AAG25940.1"  
 /db\_xref="GI:10998442"  
 /translation="MAFVBEGLIAIIVEYLILVGIWAARTKNSGAEERSEALIV  
 GGRIGLVGFMTATMVGGYINGRAVAVYVGGIAGMAQAPIGYSLILIGLTF  
 AKPMRSKGYVYMLDPFOOIYKRMGGLFPAALMGEMMAAITSAGATSVIIDV  
 WHISVITSLATITLTVGLSYVAITDVDFCIYGLMTSVFPAISHPAVDIGFT  
 AVHAKYQKPMWGLYDSEVYSWMDLFLMLGGLPQWYFQRLVSSSSATYAQVLSFL  
 AAFGLVLAIPAILIGAIASSTDMNQTAIVGLPDKTTEADMLIPIVLYQVLCPIYISF  
 FGLGAVSAVMSADSSILSASMEFARNIYOLSEFRONSADEIIVVMKITVEFGASA  
 TAMLLIKTVYGLWYLSDLIVYIYIPOLLCLVFKGNTYGAAGVYSGFLPRTGG  
 EPIYLLPFIYEPYIPDNGINYNQKPFKTLAMVSTSLNYCISYLAKYLPESGTLF  
 PKLVDFDAVVARHSEENMDKITLVKENITKIDELALVAPROSMTLSITFNKEAPLDV  
 DSSPEGGTEEDNLQ"  
 ORIGIN  
 Query Match 78.9%; Score 1375; DB 5; Length 1743;





FEATURES  
Source and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
Location/Qualifiers  
1. .1813  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2q11-13"  
/tissue\_type="hypothalamus"  
1. .1813  
/gene="CHT1"  
19. .1761  
/gene="CHT1"  
/function="sodium- and chloride-dependent reuptake of  
choline"  
/experiment="experimental evidence, no additional details  
recorded"  
/codon\_start=1  
/product="high affinity choline transporter"  
/protein\_id="CAC03717.1"  
/db\_xref="GI:9843754"  
/db\_xref="GOA:O9GZV3"  
/db\_xref="UniProtKB/TrEMBL:O9GZV3"  
/translation="MAFVVEGLIATIVVYLITLIVGIWAARTKNSGASRRSEALIV  
GARDGLVGFMTATVGGSYINGTLEAVYVPEYGLMAOAPIGISLISLGLF  
AKPMRSKYVIMLDFQOIYGRWGGLFIPALMEEMWAAIFALANTSVIIDVD  
MHSVILSALATLVLVGLYSVATVVOVLFIFGLMISVPALSHPAVADIGFT  
AHAVYQKPMGLTVDSSEVSWLDSFLMLMGLIPMOAYFQVLSSTAAQVLSFL  
AAGGLVMAIPALIGASTDMQTVGLPDPRTBEADMLIPVIOYCPVYSF  
FGLGVAASVMSADSLISASMSFARNIYOLSPRONADKEIWMWMLITVVFASA  
TAMLLITKYVGLWYSLSDIYIVIPOLLCTLPKRGNTGAVAGVYSGLFRTGCG  
EPYLYLPILYPIGYIPDDNGIYNOKFPEKTLAMVTSPLINI CISYLAKYLFESGTLF  
PKLIDVFAVVAHSEEMDKTILVNNENIKDELALVXPROSMTLSFTFMEKELDV  
DSSPEGSGTEEDNLQ"

## ORIGIN

Query Match 78.9%; Score 1375; DB 5; Length 1813;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;  
QY 1 ATGCTTTCATGTGGAAGACTGGTAGCTATATCTTCTACCTCCTTATATTTCTG 60  
DB 19 ATGCTTTCATGTGGAAGACTGATAGCTATATCTTCTACCTCCTTATATTTCTG 78  
QY 61 GTTGAATATGCGCTGCATGAAAAACAAAAACAGCGGCAACCCAGAAAGCGCATGAA 120  
DB 79 GTTGAATATGCGCTGCATGAAAAACAAAAACAGCGGCAACCCAGAAAGCGCATGAA 138  
QY 121 GCCATCATAGTCGGGGCCGTCATGTCATGTTGTTGTTGTTTTCATGACAGCC 180  
DB 139 GCCATCATAGTCGGGGCCGTCATGTCATGTTGTTGTTTTCATGACAGCC 198  
QY 181 ACCTGGTGTGAGAGAGCTACATCATGAGCAGACAGAGAGAGAGAGAGAGAGAGAG 240  
DB 199 ACCTGGTGTGAGAGAGAGCTATCATGAGCAGCAGAGAGAGAGAGAGAGAGAGAG 258  
QY 241 GGTCTAGCTTGGGCTCATGACCACTTGATATTTCTCTGAGCTTAAATTTAGGTGCTG 300  
DB 259 GGTCTAGCTTGGGCTCATGACCACTTGATATTTCTCTGAGCTTAAATTTAGGTGCTG 318  
QY 301 TTTTGTGGAACCTATGCTTCAAGGATATGATGATGATGATGATGATGATGATGATGAT 360  
DB 319 TTTTGTGGAACCTATGCTTCAAGGATATGATGATGATGATGATGATGATGATGATGAT 378  
QY 361 ATCTATGAAAGCGAGTGGGTGGCTCTTCACTCCCTGACAGAGAGAGAGAGAGAGAG 420  
DB 379 ATCTATGAAAGCGAGTGGGTGGCTCTTCACTCCCTGACAGAGAGAGAGAGAGAGAG 438  
QY 421 TGGGCTGAGCAATTTTCTCTGCAATTTAGGGGCAACCATGAGCGTATCATTTAGTGTGAT 480  
DB 439 TGGGCTGAGCAATTTTCTCTGCAATTTAGGGGCAACCATGAGCGTATCATTTAGTGTGAT 498  
QY 481 GTGAACATATGGGTATGTTCTCTGCACTCATTTGCAATTTTATACCTATGAGGTGAG 540

DB 499 ATGCAATTTCTGATCATATCTCTGCACTATTCGCACTCTGTACACACTGATGGAGG 558  
QY 541 CTCTACTGTGAGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 559 CTCTACTGTGAGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618  
QY 601 ATCACTGTCCCTTTTGGCCCTGTCAATCTGCACTGACCGACATGATGATGATGATGAT 660  
DB 619 ATCACTGTCCCTTTTGGCCCTGTCAATCTGCACTGACCGACATGATGATGATGATGAT 678  
QY 661 CATGCTAATATACAGAGTCCCTGTGAGAACATGATGATGATGATGATGATGATGATGAT 720  
DB 679 CATGCTAATATACAGAGTCCCTGTGAGAACATGATGATGATGATGATGATGATGATGAT 738  
QY 721 CTCTAATATTTCTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB 739 CTCTAATATTTCTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 798  
QY 781 GTCTCTCTCATCTCTGAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 799 GTCTCTCTCTCTCTCTGAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 858  
QY 841 TGCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 859 TGCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918  
QY 901 AACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 919 AACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978  
QY 961 CCGATCGTTCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 979 CCGATCGTTCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038  
QY 1021 TCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
DB 1039 TCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098  
QY 1081 CGGAATATCTACAGCTTCTCTGAGCAAAATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1099 CGGAATATCTACAGCTTCTCTGAGCAAAATGATGATGATGATGATGATGATGATGATGAT 1158  
QY 1141 ATGAGATCATCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200  
DB 1159 ATGAGATCATCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1218  
QY 1201 ACTGTGTATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1219 ACTGTGTATGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278  
QY 1261 CTGCTCTGTGATCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1279 CTGCTCTGTGATCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1338  
QY 1321 TTTGAGCATTTCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
DB 1339 TTTGAGCATTTCTGAGAAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398  
QY 1381 TTTTACCTGCTTATTTATCTGACAAAGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1399 TTTTACCTGCTTATTTATCTGACAAAGATGATGATGATGATGATGATGATGATGATGAT 1458  
QY 1441 ACTCTCTCATAGTATCTCATTTCTTATGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1459 ACTCTCTCATAGTATCTCATTTCTTATGATGATGATGATGATGATGATGATGATGATGAT 1518  
QY 1501 CTATTTGAAAGTGAACCTTGTCTTCAAAATTAATGATGATGATGATGATGATGATGAT 1560  
DB 1519 CTATTTGAAAGTGAACCTTGTCTTCAAAATTAATGATGATGATGATGATGATGATGAT 1578  
QY 1561 CACAGTGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620



Db 1579 CACAGTGAAGAAACATGATAGCAATTCTTGCAAAATGAAATATTAATTAGAT 1638  
QY 1621 GAACCTTGACACTGTGAAACCTCGGACAGAGCCTAACCTTCAGTTCAACTTTCACCAATAG 1680  
Db 1639 GAACCTTGACACTGTGAGCCACGACAGACATGACCTCAGCTCAACTTTCACCAATMAA 1698  
QY 1681 GAGGCCCTCTGATGTTGATTCAGTCCGAGAGGGGCTGGGACTGAGATAATTACAA 1740  
Db 1699 GAGGCCCTCTGATGTTGATTCAGTCCAGAGGGGCTGGGACTGAGATAATTACAG 1758  
QY 1741 TGA 1743  
Db 1759 TGA 1761

Search completed: July 11, 2006, 02:51:36  
Job time : 9697 secs

**This Page Blank (uspio)**

Result No.	Score	Query Match	Length	DB	ID	Description
1	2993	100.0	1743	6	AF276872	Mus musculus
2	2973	99.3	4715	6	BC065089	Mus musculus
3	2968	99.2	1743	2	BD012720	High-altit

	4	2968	99.2	1743	2	E49872	High-affini
	5	2967	99.1	4938	2	AX08043	Sequence
	6	2958	98.8	1743	6	MMU041467	Mus muscu
	7	2956	98.8	1743	2	BD012718	High-affin
	8	2956	98.8	1743	2	E49870	High-affin
	9	2956	98.4	4904	6	AB030947	Rattus no
	10	2795	93.4	1743	2	BD012719	Rattus no
	11	2795	93.4	1743	2	C0722844	High-affi
	12	2795	93.4	1743	2	E49871	Sequence
	13	2795	93.4	1743	2	AR268949	High-affin
	14	2795	93.4	1743	5	AF276871	Sequence
	15	2795	93.4	1813	5	HSA0401466	Human sapi
	16	2795	93.4	2583	5	BC111525	Human sapi
	17	2795	93.4	5158	5	AB043997	Human sapi
	18	2695	90.0	2541	5	BC111524	Human sapi
	19	2680	89.5	2112	14	BC112794	Bos tauru
	20	2226	74.4	2528	11	TMA420808	Torpedo m
	21	1696	56.7	166000	11	AC174718	Gasterost
	22	1694	56.6	1132	11	GGAS11267	Gallus ga
	23	1542.5	51.5	3255	13	AY047521	Drosophi
	24	1517	50.7	3226	13	AY011119	Limulus p
	25	1489.5	49.8	2049	13	AY629593	Trichoplu
	26	1444.5	48.3	1731	2	BD012717	High-affi
	27	1444.5	48.3	1731	2	E49869	High-affin
	28	1444.5	48.3	1985	13	AB030946	Caenorhab
	c 29	1413	47.2	4223	5	CG613835	Sequence
	30	1413	47.2	186989	13	AC007812	Drosophi
	31	1413	47.2	189116	13	AC009395	Drosophi
	32	1413	47.2	25619	13	AE003723	Drosophi
	33	1405	46.9	17259	2	C0613836	Sequence
	c 34	1386	46.3	140156	12	AC017381	Drosophi
	c 35	1200.5	40.1	155210	6	CT102873	Mus muscu
	c 36	1200.5	40.1	257445	12	CT571242	Mus muscu
	c 37	1152.5	38.5	40893	13	CBR445E19	Rattus no
	38	1136	38.0	232792	12	AC106657	Caenorhab
	c 39	1105.5	36.9	39908	13	CEC48D1	Rattus no
	40	1103	36.9	110000	15	AE017933_31	Caenorhabd
	41	1103	36.9	110000	15	CP000022_31	Continuati
	42	1099.5	36.7	1461	2	AX432086	Sequence
	c 43	1097	36.7	159128	12	AC182332	Pan trogl
	c 44	1096	36.6	190043	5	AC009963	Homo sapi
	c 45	1094.5	36.6	330724	12	CEY67H2	Alouatta

## JOURNAL

Submitted (28-FEB-2001) Department of Pharmacology and Center for Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at Pierce, Nashville, TN 37232-6420, USA

## REMARK

Sequence update by submitter

## COMMENT

## FEATURES

On Feb 28, 2001 this sequence version replaced gi:11527247.

## Location/Qualifiers

1..1743  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..1743  
/codon\_start=1  
/product="sodium and chloride-dependent high-affinity choline transporter"  
/protein\_id="AAG36945.2"  
/db\_xref="gi:11527247"  
/translation="MPFVEGLVATILFYLILFLVGIWAAMTKNSGNPSESEALIVGGDILGVGFTMTATWVGGSYINGTAEAVYPGCGIAMAAPIGYLSILIGLFP AKPRMSKSYVTMDPFOKIYGRMGSLFIPALMEEMRAAIPESALGATISVITD VNIISVALIAILYTLVGLYSVAITDVQLCIFIGLWISVPALSHPATDIGFT AVHAKYQSPWLGITISVEVYTWLDFLMLGIPMOYFORWISSSATAQVVISFL AAFGLVVALPAICIGAIIGASIDMNTAYGYDPKTKESADMLPILVQYICPVVISF FGLGAVSAAMNSADSSILISASMPFARNIYOLSPRONSDKEIWMWMTITVFCGASA TAMALITKTVYGLWYLSDDIVYIIIFPOLICVLFIKNTYGAVAGYIEGLRITGG EPIYLQPLIFPIGYSDKNGIYNORPFKTIISMTSPPTNI CYSYIAKYLPESGTLR FKLVFDVAVHSHSENMKTIIVRENIKIMELAPVAKROSITLSTFTNKEALLDV DSSPESGSEEDMLQ"

## CDS

## ORIGIN

## Alignment Scores:

Pred. No.: 0  
Score: 2993.00  
Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 100.0%

Length: 1743  
Matches: 560  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-724-806-4 (1-580) x AF276672 (1-1743)

1 MetProphHisValGluGlyLeuValAlaIleLeuPheTyrLeuLeuIlePheLeu 20  
1 ATGCTTTCCATGTGGAAGGACTGAGTATTAATCTCTTCTACCTCTTATATTTCG 60  
21 ValGlyIleTrrPalAlaIleTrrPlyThrIleAsnSerGlyAsnProGluIleArgSerGlu 40  
61 GTTGGAAATAGGCGTCATGAGAAACCAAAACAGGGCAACCCAGAAAGCGCAGTGA 120  
41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60  
121 GCCATCATATGCGGGGCCGCTGACATTTGTTGTTGGTGGTTTTCACATGACAGCC 180  
61 ThrTrrPalGlyIleGlyIleTrrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
181 ACCTGGGTTGGAGGAGGCTACATCATGAGCAGCAGAGCAGTGAATGGCCAGGTTGT 240  
81 GlyLeuAlaIleTrrPalAlaHisAlaProIleGlyTrrSerLeuSerLeuIleLeuGlyIleLeu 100  
241 GGTCTAGCTTGGGCTCATGACCCATTTGATTTCTGAGTCAATTTTAACTGAGTGTG 300  
101 PhePheAlaIleYsProMetArgSerIleYsGlyTrrValThrMetLeuAspProPheYsGln 120  
301 TTTTTCGGAAACCTATGCTGTTCCAGGATATGTGACTATGTAGACCATTCACAAACG 360  
121 IleTrrGlyIleYsArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleLeuPhe 140  
361 ATCTATGAAAGCGCATGGGTGGCTGCTTCATCTGCACTGATGGAGAGATGTTTC 420  
141 TrrPalAlaAlaIlePheSerAlaLeuGlyAlaTrrIleSerValIleIleLeuSpsValAlaP 160  
421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGGCCACCATCAGCGTATCTTAATGAGAT 480  
161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyIle 180

481 GTGAACATATGCGTCATGTTCTCTGCACTCATTTGCCATTTCTTATACCTTACGTCGCG 540  
181 LeuTrrSerValAlaIleTrrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTrrP 200  
541 CTCTACTCTGAGGAGATATACGATAGTGTTCAGGATATCTGCAATTTTATAGACTGTGG 600  
201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
601 ATCATGTCCTCTTGTGGCCCTGCTCATCTTCACAGCAGCAGCATCGATTCACACCTGTG 660  
221 HisAlaIleYsTrrGlnSerProTrrPheGlyIleThrIleGlnSerValGluValTyrThrTrrP 240  
661 CATGCTTAATATACAGAGTCCTGCTGGGAACATTTGAATGATTTGAAGTCTACACCTGG 720  
241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTrrGlnAlaTyrPheGlnArg 260  
721 CTGTATTAATTTTCTGTATTTATGATGCTGGGTGAATCCATGAGCAAGCTTACACTTCAGAG 780  
261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280  
781 GTCCCTCTCTTATCTTCAGCCACCATATGCTCAGGTAATGCTCTTCTGGCAGCTTTTGG 840  
281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrrP 300  
841 TCCCTGGTGAATGCTCTTACCCGCAATATGATAGAGCTATTGGAGCTTCCACAGACTGG 900  
301 AsnGlnThrAlaTyrGlyIleTrrProAspProIleYsGlnGluAlaAspMetIleLeu 320  
901 AACCAAGCTGCTTACGGGTATTCAGATCCCAAGACTAAGAGAGAGCAAGCATGATTC 960  
321 ProIleValLeuGlnTrrLeuCysProAlaTrrIleSerPhePheGlyLeuGlyAlaVal 340  
961 CGGATGCTTTCGACATACCTCTGCTTGTGACATCTCTTCTTGGGCTTGGTCTGT 1020  
341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
1021 TCACTGCTGTCATCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
361 ArgAsnIleTrrGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValTrrPal 380  
1081 CGGAATATCTACACACTTCTCTTCCAGCAAAATGATCAGACAAAGAAATGTGTGGGCTC 1140  
381 MetArgIleTrrValLeuValPheGlyAlaSerAlaTrrAlaMetAlaLeuLeuThrIle 400  
1141 ATGAGGATCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1200  
401 ThrValTyrGlyLeuTrrPlyLeuSerSerAspLeuValTrrIleIlePheProGln 420  
1201 ACTGCTATGGGCTTGGTATCTGATCCTGATCCTTGTCTATCATCATCTTCCACAG 1260  
421 LeuLeuCysValLeuPheIleIleYsGlyThrAsnThrTyrGlyAlaValAlaGlyTrrIle 440  
1261 CTGCTCTGTGATCTTCTTATCAAAAGAAACCAACTTATGGGGCAGTGTCTGATATAT 1320  
441 PheGlyLeuPheLeuArgIleIleThrGlyIleGluProIleTrrLeuGlnProLeuIle 460  
1321 TTTGACATATCTCTGAAATTAATGAGAGAGCCATATCTATCTTGAAGCTTAAATC 1380  
461 PheTrrProGlyTrrTrrSerAspIleYsAsnGlyIleTrrAsnGlnArgPheProPheYs 480  
1381 TTCTACCTCGGTATTTACTCTGACAGAAATGTATATACATCAGAGGTTCCATTAA 1440  
481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrrLeuAlaIleYsTrr 500  
1441 ACTCTCTCAAGTATCACTCATCTTCTTACCAACTTGTGTTCTTATCTAGCCCAAGAT 1500  
501 LeuPheGlnSerGlyThrLeuProProIleYsAspValPheAspAlaValAlaIleArg 520  
1501 CTATTTGAAAGGAAACCTTGTCTCAAAATTAAGATTAATTAAGTCTGTTGTCCAAAG 1560  
521 HisSerGlnGluAspMetAspIleThrIleLeuValArgAsnGluAsnIleYsLeuAsn 540  
1561 CACACTGAAGAGAAATGAGCAAGCACTTCTAGTCABAATGAATAATTAATTAAT 1620

QY	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS
54	GlueuulnlaPrcovLlySPrcvAGlSerleuThrluSerSerThrPhaThAsnLys	560						
Db	1621	GAACCTTCGACCTGTGGAAACCTCGGACAGACCTTAACCTCAGTTACCTTACCAATTAAG	1680					
QY	561	Glua1a1eulnleuAepValjAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln	580					
Db	1681	GAGGCGCTCTTGAATGTTGATTCCAGTCCGAGGAGGTCTGGAGACTGAAGATTAATTACA	1740					
RESULT 2	BC065089							
LOCUS	BC065089							
DEFINITION	Mus musculus solute carrier family 5 (choline transporter), member 7, mRNA (cDNA clone MGC:86023 IMAGE:6851099), complete cds.							
ACCESSION	BC065089							
VERSION	BC065089.1							
KEYWORDS	GI:40787811							
ORGANISM	Mus musculus (house mouse)							
REFERENCE	1 (bases 1 to 4715)							
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S., Caranci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Holty, S.W., Villalton, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schercher, A., Schein, J.E., Jones, S.J., and Marra, M.A.							
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences							
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.							
PUBMED	12477932							
REFERENCE	2 (bases 1 to 4715)							
AUTHORS	Strausberg, R.							
TITLE	Direct Submission							
JOURNAL	Submitted (02-JAN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA							
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov							
COMMENT	Contact: MGC help desk Email: cgaabs-remail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant. Web site: http://genome.uiowa.edu Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu Ronald, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Smith, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.							

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov  
Series: Place: Row: Column: 0  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11528509  
Location/Qualifiers

[illegible]

```

Db      387 ATCTATGAAAGCGATGGGCTGCTCTTCATCCCTGCACGTAGTGGAGAGATGTTTC 446
Qy      141 TTPALAAALALEPHESEALALEUGLYALATHRIILESEVALIILELEASVALASP 160
Db      447 TGGGCTGCAGCAATTTTCTCTGCATTAAGGGGCCACATCAGCGTATCTTATGATGGAT 506
Qy      161 VALANILIESEVALIILEVALSERALALEUILEALILEUITYRTHLEUVALGLYGLY 180
Db      507 GTGAACATATCGGTATGTCCTCTGCACATTCGACATTCCTTAATACCTTACGTGGGTGG 566
Qy      181 LEUITYRSEVALALETYRTHASPYALVALGINLEUPHECYEIIPEHEILEGYLEUSTRP 200
Db      567 CTCTACTCTGGGCATATCTCATGTGTCCAGCTATTTCTGCATTTTATAGACCTGG 626
Qy      201 ILESEVALPROPEALALEUSERHSIPROALAVALTHASPIILEGYPHETHRALAVAI 220
Db      627 ATCAGTGTCCTTTGGCCCTGCACATCTGCAGATCCAGCATGGATTCACAGCTGG 686
Qy      221 HISALALYTYRGLINSEPROTRPLEUGLYTHRILEGUSEVALGLUVALTYRTHTRP 240
Db      687 CAVGTAAATATACCAAGTCCCTGGCTGGGAAACCATTTGAATCACTTACACCTGG 746
Qy      241 LEUASPANPHELEULEUMETLEUGLYGLYILEPROTRPGALALATYRPHGLINARG 260
Db      747 CTGTGTAATTTTCTGTATGATGCTGGGTGGAAATCCATGGCAACCTTACCTTCAGAG 806
Qy      261 VALLEUSERSESESERIALATHRYALAGINVALLEUSERPHLEUVALAIAPEHGLY 280
Db      807 GTCCCTCTTCATCCCTCAGCCACCATATGCTCAGTACTGCTCTCCCTGGCAGCTTTGGG 866
Qy      281 CYSLEUVALMETALEUPROBALILECYRIIEGLYALALEGLYALASERPHTRSPTRP 300
Db      867 TGCCGTGGTGAATGCTTACCCCAATGACATAGAGAGCTATTGGAGCTTCCACAGCTGG 926
Qy      301 AENGINTRALATYRGLYTYRPROASPPLYSTHLYSGULUALASPEHILELEU 320
Db      927 AACCAAGACTGCTTACGGGTATCCAGATCCCAAGATCAAGAGAGACACATGATTC 986
Qy      321 PROILEVALLEUGINTYRLEUCYSPROVALIYRIILESEPHPEHGLYLEUGLYALAVAI 340
Db      987 CCGATCGTTCGACATCTGCTGCGCTGTGATCTCTTGGGCTGTGGTGTGCTGT 1046
Qy      341 SERIALAVALMERSESERIALASPSERILELEUSERALASERSETHPEALIA 360
Db      1047 TCAGCTGCTGATATGCTCCAGCTGATCTGTCATCTGTCGAGATTCATAGTTTGT 1106
Qy      361 ARGANILETYRGLINLEUSERPHARGINANALASERASPLYSGULILEVALI 380
Db      1107 CGGAATATCTACCACTTCTTCAGACAAATGATCAGACAGAAATGTGTGGGTTC 1166
Qy      381 METARGILETHVALLEUVALPHEGLYALASERIALATHRALMETALALEULEUTHLYS 400
Db      1167 ATGAGGATCATCTGCTGTGTGTCGAGCATGCAACAGCCATGGCTTGTGTCAGGAAG 1226
Qy      401 THRVALITYRGLYLEUTRYRLEUSERSEASPLLEUVALTYRIILELEIPEHEPROGIN 420
Db      1227 ACTGTGATGGCTCTGTGACTGAGCTTGACTTGCTTGAATCATCTTCCACAG 1286
Qy      421 LEULEUCYVALLEUPHEILEYSGLYTHASNTHYRGLYALAVAIAGLYTYRI 440
Db      1287 CTGCTCTGTGATCTTTCATCAAGGAACCAACTTAAGGGGCACTTCTGTTATATT 1346
Qy      441 PHEGLYLEUPHELEUARGILETHRGLYGLYIUPROTYRLEUITYRLEUGINPROLEU 460
Db      1347 TTTGACATATTCCTAGATTAATCTGGAGAGAGCCATATCTTACTTTCAGCCCTTAATC 1406
Qy      461 PHEITYRPROGLITYRYSERASPLYASNGLYILETYRASNGLINARPHPEPROPEHYS 480
Db      1407 TTCTACCTGCTTATCTCTACAGAAATGGTATATCAATCAGAGGTTCCCATTTAA 1466
Qy      481 THRLEUSERMETVALTHSERPHEPETHASNILECYSEVALSEITYRLEUVALIAYSYR 500
Db      1467 ACTCTCTCATGGTATCTCATTTCTTACCAACATTTGTTCTTATCTACCAAGTAT 1526

```

```

Qy      501 LEUPHEGLUSERGLYTHLEUPROPROLYSELEUASPVALPHEASPVALVALIAARG 520
Db      1527 CTATTGAAAGTGAACCTTGCTCCAAATTAATGATTAATTAATGATGCTGTGTCCAGAG 1586
Qy      521 HISSERGULUGIARMEASPLYSTHRIILEUVALAARGANGULASNNILEYSELEUN 540
Db      1587 CACAGTGAAGAGAAATGACAAAGCAACTTCTAGTCAGAAATGAAATATCAATTAAT 1646
Qy      541 GULEUALAIPROVALIYSPROARGINSEURLEUTHLEUSERSETHPHETHASNLYS 560
Db      1647 GAACCTTGACCTGTGAAGAACTCGGAGAGCTTAAACCTCAGTTCACCTTACCAATTAAG 1706
Qy      561 GUALALEULEUASPVALASPSERSEPROGLUGIYSERGLYTHYTHGLUASPANLEUGIN 580
Db      1707 GAGGCTCTCTGATGATGATTCCTCAGTCGAGAGGGGCTCGGACATGAAATTAATTAACA 1766

RESULT 3
BD012720
LOCUS      BD012720          1743 bp      DNA      linear      PAT 02-AUG-2002
DEFINITION High-affinity choline transporter.
ACCESSION  BD012720.1 GI:22092909
VERSION    BD012720.1
KEYWORDS   WO 0116315-A/4.
SOURCE
ORGANISM   Mus musculus (house mouse)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1743)
AUTHORS    Haga, T. and Okuda, T.
TITLE      High-affinity choline transporter
PATENT     Patent: WO 0116315-A 4 08-MAR-2001;
JOURNAL    JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA
COMMENT
OS          Mus musculus (mouse)
PN          WO 0116315-A/4
PD          08-MAR-2001
PF          18-AUG-2000 WO 2000JP005545
PR          27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI
TATSUYA HAGA, TAKASHI OKUDA
PC          C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC
            A61K38/17,
            PC A61K45/00, A61P25/28, G01N33/53, A01K67/027
CC          CC
FH          FH
FT          CDS
FEATURES
source      1..1743
             /organism="Mus musculus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      1743
Percent Similarity: 99.5%      Matches:      576
Best Local Similarity: 99.3%      Mismatches: 1
Query Match:    99.2%      Indels:      3
DB:              2          Gaps:      0
US-10-724-806-4 (1-580) x BD012720 (1-1743)
Qy      1 METPROPEHISVALGLUGLYLEUVALAIAIILELEUPHEITYRLEUENILEPHELEU 20
Db      1 ATGTCTTCCACGTGAGAGACTGGTAGCTATTAATCTCTCTACTTCTATATTTCTG 60
Qy      21 VALGLYILETPALALATITRPLYSRTHLYSANSERGLYASNPROGLUGIARSGEGLU 40
Db      61 GTTGAATATGGCTGCAATGAAACCAAAAACAGGGCAACCCGAAAGACACGTGA 120
Qy      41 AIAIILEIIVALEGLYARASPIILEGLYLEUENVALGLYGLYPHETHMETETHALA 60

```

Db 121 GCCATCATAGTCGGGGCCGTGACATTGGTTGGTTGGTTTATCCATGACAGCC 180  
 QY ThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
 Db 181 ACCGTGGGTGGAGAGGCTACATCAATGGAGACAGAGCATGTATGGGCCAGGTTGT 240  
 QY GlyLeuAlaTrpAlaHisAlaProIleGlyTyrSerLeuSerLeuIleGlyGlyLeu 100  
 Db 241 GGTCTAGCTTGGGCTCAGGACCCATTGGATATCTCTGAGTCTTATTTAGTGGTCTG 300  
 QY PhePheAlaLysProMetArgSerIleGlyTyrValThrMetLeuAspProPheLysGln 120  
 Db 301 TTTTTCGCAAACTGTGCGTTCCAGGAGATAGTCACTATGTTAAGCCATTTCACACG 360  
 QY IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
 Db 361 ATCTATGAAAGGAGGAGGAGTGGGCTGCTTCATCCCTGCACTGATGGAGAGATGTC 420  
 QY TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db 421 TGGGCTGCAGCAATTTCTCTGATTAAGGGCCACCATCAGCGTATCATGTATGTGAT 480  
 QY ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
 Db 481 GTGAACATATCGGTCAATTGCTCTGCACTCATTCGCAATTCCTTATACCTAGTGGTGG 540  
 QY LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
 Db 541 CTCTACTCTGGGCAATATCACTGATGTGTCACACTATTCGCAATTTTATAGGACGTGG 600  
 QY IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 Db 601 ATCAGTGTCCCTTTTCCCTGCTGACATCTCGACGTCAACGACATCGGATTCACAGCTGG 660  
 QY HisAlaLysTyrGlnSerProTyrIleGlnSerValGluValTyrThrTyr 240  
 Db 661 CATGCTAATATACAGAGTCCCTGGCTGGGNAACATTGAATCAGTTGAAGCTCACCTGG 720  
 QY LeuAspAspPheLeuLeuMetLeuGlyGlyIleProThrGlnAlaTyrPheGlnArg 260  
 Db 721 CTGTATATTTTCTGTATGATGCTGGGGTGGAGATCCCATGGACCTTCCACAGAG 780  
 QY ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
 Db 781 GTCTCTCTTCTATCTTCAGGCACTTATGCTCAGGTACTGCTTCTCTGGAGCTTTGGG 840  
 QY CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
 Db 841 TGCCTGTGATGGCTTACCCCGCATATGATGATGAGCTATTGGAGCTTCCACAGACTGG 900  
 QY AsnGlnThrAlaTyrGlyTyrProAspProLysThrIleGlnGluAlaAspMetIleLeu 320  
 Db 901 AACCAAGACGCTTACGCGGTATCCAGATCCCAAGACATTAAGAGAGAGACATGATTTCTC 960  
 QY ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
 Db 961 CCGATGCTTCTGAGTACCTCTGCTGCTGCTATCATCTCTTCTTGGGCTTGGTGTCTGT 1020  
 QY SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
 Db 1021 TCAGCTGCTGTCAATGCTCTCAGCTGACTCGTCATCTCTGCGCGAGTTCTATGTTGCT 1080  
 QY ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerSerLysGlyIleAlaTrpVal 380  
 Db 1081 CGGAATATCTACACAGCTTCTCTCAGCAAAATGATCAACAAAGGAATATGTGTGGTTC 1140  
 QY MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLys 400  
 Db 1141 ATGAGATCACTGTGCTGTGTGTGCGAGCATCTGCAACACCATGGCTTGTGACGAG 1200  
 QY ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
 Db 1201 ACTGTATATGGGCTGTGTAACGAGCTGACCTTGTCTTACATCATCATCTTCCACAG 1260

QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
 Db 1261 CTGCTCTGTCTACTCTTCAACAGAACCACTTATGGGGCACTTGTGTTATTT 1320  
 QY PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460  
 Db 1321 TTTGACATATTCCTGAGATTAATCTGGAGAGAGCCATATCTATCACTTGCAAGCCCTTAATC 1380  
 QY PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
 Db 1381 TTTCTACCTGGTTATTACTGACAGAAATGATATACAAATCAGAGTTCCCATTTAA 1440  
 QY ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
 Db 1441 ACTCTTCATGTTTACCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAAAGTAT 1500  
 QY LeuPheGlnSerGlyThrLeuProProlLysLeuAspValPheAspAlaValAlaArg 520  
 Db 1501 CTATTGAAAGTGAACCTTGCTCCAAATTAAGATGTATGATCTGTGCGCAAGG 1560  
 QY HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
 Db 1561 CACAGTGAAGAGAACTGACAAAGACCATTTAGTCAAGAAATGAATAATCAATTAAT 1620  
 QY GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
 Db 1621 GAACTTGCACTGTGAACCTCGGACAGAGCTTACCTCAGTTCACTTACCAATTAAG 1680  
 QY GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyTyrGluAspAsnLeuGln 580  
 Db 1681 GAGGCCCTCTTGATGTATGTTATTCAGTCCGAGAGGGGTCTGGAGACTGAAGATTAATCAAA 1740

## RESULT 4

E49872 1743 bp DNA linear PAT 27-AUG-2002  
 LOCUS High-affinity choline transporter.  
 DEFINITION E49872  
 E49872.1 GI:22554903  
 VERSION JP 2001136976-A/4.  
 KEYWORDS Mus sp.  
 SOURCE  
 ORGANISM Mus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1743)  
 AUTHORS Haga,T. and Okuda,T.  
 TITLE High-affinity choline transporter  
 JOURNAL Patent: JP 2001136976-A 4 22-MAY-2001;  
 SCIENCE & TECH AGENCY

COMMENT OS Mus sp. (mouse)  
 FN JP 2001136976-A/4  
 PD 22-MAY-2001  
 PE 27-DEC-1999 JP 1999368991  
 PI TATSUYA HAGA, TARASHI OKUDA  
 PC C12N15/09,A01K67/027,A61K38/00,C07K14/47,C07K16/18,C07K19/00,  
 PC C12N15/10,  
 PC C12P21/02,C12P21/08,C12Q1/00,C12N15/00,A61K37/02,C12N5/00 CC  
 FH Key Location/Qualifiers  
 FT CDS Location/Qualifiers  
 1..1743  
 1..1743  
 /organism="Mus sp."  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10095"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 1743  
 Score: 2968.00 Matches: 576  
 Percent Similarity: 99.5% Conservative: 1  
 Best Local Similarity: 99.3% Mismatches: 3  
 Query Match: 99.2% Indels: 0

DB: 2 Gaps: 0  
US-10-724-806-4 (1-580) x E49872 (1-1743)  
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleuPheThrLeuLeuIlePheLeu 20  
DB 1 ATGTCTTCCAGTGAAGAGCTGAGTATTAATCTCTTCTTCACTCTTAATTTCTG 60  
QY 21 ValGlyIleTPAlaAlaTPLeuThrLySaenSerGlyAsnProGluGluArgSerGlu 40  
DB 61 GTTGGAAATATGGGCTGCATGGAACCAAAAACAGGGCAACCCAGAAAGACAGGTGA 120  
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
DB 121 GCCATCAAGATGGGGGCGGTGACATGGTTGTGTGTGTGTGTATTCATGACAGCC 180  
QY 61 ThrTPValAGlyGlyGlyTyrIleLeuGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
DB 181 ACCTGGGTTGGAGAGGCTTACATCAATGGGACAGCAAGAGCTGATGGCCAGGTTGT 240  
QY 81 GlyLeuAlaTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
DB 241 GGTTAGCTTGGGCTCAGGCAACCATTTGGATATTTCTGAGTCTAATTTTAGTGGTCTG 300  
QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120  
DB 301 TTTTGTGGAAACCTAATGGTTCCAAAGGATATGTGACTATGTAGACCATTTTCACACG 360  
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
DB 361 ATCTATGGAAAGCGATGGGTGGGCTGCTCTTCAATCCCTGCACCTGATGGAGAGATGTC 420  
QY 141 TPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
DB 421 TGGGCTGACAGCAATTTTCTCTGCATTAAGGGCCACCATCAGCGTATCTATGATGGAT 480  
QY 161 ValAlaIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
DB 481 GTGAACATATCGGTATGTCTCTGCACATTCGCAATTCCTTAATACCTCAATGGGTGG 540  
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
DB 541 CTCTACTCTGTGGCATATCTGATGTGTCCAGCTATTTCTGCATTTTATAGACTGTGG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
DB 601 ATCAGTGTCCCTTTGGCCCTGTCAATCTCTGACATCCGACATCGGATTCACAGCTGTG 660  
QY 221 HisAlaLysTyrGlnSerProTPLeuGlyThrIleGlnSerValGluValTyrThrTyr 240  
DB 661 CATGTCTAAATATACCAAGTCCCTGGCTGGGAGACCATTTGAATCGATTCAACCTGG 720  
QY 241 LeuAspAsnPheLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260  
DB 721 CTGTGATATTTTCTGTATATGATGCTGGGTGGAATCCCATGGCAACCTTATCCAGAG 780  
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
DB 781 GTCCCTCTCTTCACTCTCAGCCACCTATGCTCAGTACTGTCTCTCTGACACTTTGGG 840  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
DB 841 TGCCGTGGTATGGCTTCAACCCCATATGCAATAGGAGCTAATGGAGCTTCCACAGCTGG 900  
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlyGluAlaAspMetIleLeu 320  
DB 901 AACCAAGACTGCTTACGGGATCCAGATCCCAAGACATRAAGAGACACATGATTTCTC 960  
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
DB 961 CCGATCGTTCTGACATCTGCGCTGTGTACATCTCTCTTGGGCTGTGTGTCTGTT 1020  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360

DB 1021 TCAGCTGCTGCATAGTCCCTCAGCTGACTGTCACCTCGTCGCGCAGTTCTATGTTTCT 1080  
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTPVal 380  
DB 1081 CGGAATATCTACAGCTTCTCTTCAGACMAAATGATATGACACAGAAATGTGTGGGTG 1140  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
DB 1141 ATGAGATACATGTCCTGTGTTCGAGACATCTGCACAGCCATGCTTGTCTGACGAAG 1200  
QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
DB 1201 ACTGTATAGGCTCTGTGTACCTGAGCTGTGACTGTGCTATCATCATCTTCCACAG 1260  
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
DB 1261 CTGCTCTGTGTCCTTCTTATCAAGAGAACCAACACTTATGGGCGAGTGTCTATAT 1320  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrLeuTyrLeuGlnProLeuIle 460  
DB 1321 TTTGACTATTCCTGAGAAATTACTGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1380  
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
DB 1381 TTCTACCTGTGTATTACTCTGACCAAGATGTATATCAATCAAGAGTTCCCATTTAAA 1440  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
DB 1441 ACTCTCTCAATGTTACTCTCATTTCTTACCAACATTTGTGTCTTATCTAGCCCAAGAT 1500  
QY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520  
DB 1501 CTATTTGAAAGTGAACCTTCCTCAAAATTAAGATGATTTGATGAGCTGTGTCCAGAG 1560  
QY 521 HisSerGlnGluAsnMetAspLysThrIleLeuValaGlnGluAsnIleLysLeuAsn 540  
DB 1561 CACAGTGAAGAGAAATGACACACATTCAGTCAAGAAATGAAATATCAAAATTAAT 1620  
QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
DB 1621 GAACCTTGACCTGTGAACCTCGGACAGCCTAACCTCACTTCAACCTTAAG 1680  
QY 561 GluAlaLeuLeuAspValAspSerSerProGlnLysSerGlyThrGluAspAsnLeuGln 580  
DB 1681 GAGGCCCTCTTGTATGATTCACAGTCCGAGGGGCTGTGGACTGAAGATTAATTACAA 1740  
RESULT 5  
AX080443 LOCUS 4938 bp DNA linear PAT 22-FEB-2001  
DEFINITION Sequence 1 from Patent WO0078950.  
ACCESSION AX080443  
VERSION AX080443.1 GI:13159872  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 Sierzege, M. and Albrandt, K.  
Differentially expressed genes in the adipocytes of obese mice  
Patent: WO 0078950-A 1 28-DEC-2000;  
AMYLIN PHARMACEUTICALS, INC. (US)  
FEATURES  
source 1..4938  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"  
/note="P4P6B1"  
ORIGIN  
Alignment Scores:



Pred. No.: 0 Length: 4938  
 Score: 2967.00 Matches: 576  
 Percent Similarity: 99.5% Conservative: 1  
 Best Local Similarity: 99.3% Mismatches: 3  
 Query Match: 99.1% Indels: 0  
 DB: 2 Gaps: 0

US-10-724-806-4 (1-580) X AX080443 (1-4938)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheThyLeuLeuIlePheLeu 20  
 Db AGCTCTTCCAGTAAAGAGACGTAGCTATTATCTCTTCACTCTTAACTCACTTACCTG 306  
 QY 21 ValGlyIleTyrAlaIleTyrPheThyAsnSerGlyAsnProGluIleArgSerGlu 40  
 Db GTTGGAAATATGGGCTGCATGAAACCAAAACAGGGCAACCCAGAAAGCCAGTGA 366  
 QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThra 60  
 Db GCATCATAGTCGGGGCCGTGACATTGGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 426  
 QY 61 ThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGly 80  
 Db ACCGTGGGTGGAGAGAGCTACATCATGGAGACAGAGCACTGATGGCCAGGTTGT 486  
 QY 81 GlyLeuAlaTyrAlaIleHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGly 100  
 Db GGTCTAGCTTGGGCTCAGGACCCATTTGATTTCTTCTGATTTATTTAGTGGTCTG 546  
 QY 101 PhePheAlaLeuSerMetArgSerGlyTyrValThrMetLeuAspPropheGly 120  
 Db TTTTTCGCAAACTATGCGTTCCAGGGATATGATGATGATGATGATGATGATGATG 606  
 QY 121 IleTyrGlyLysArgMetGlyGlyLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
 Db ACTATGAGAAAGCGCATGGGTGGCGCTTCTCATCTGCACTGATGGAGAGAGTTTC 666  
 QY 141 TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db TGGGCTGCAGCAATTTTCTCTGATTAAGGGCCACATCAGCTGATTCATTGATGGAT 726  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
 Db GGAACATATCGGCTCATGCTCTGCACATCATGTCATTCTTATACCTTACCTGAGTGG 786  
 QY 181 LeuTyrSerValAlaTyrThrAspValValGluPheCysIlePheIleGlyLeuTyr 200  
 Db CTCTACTCTGTGGCATATCTGATGTGTCCACCTATTCTGCATTTTATAGGACTGTGG 846  
 QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 Db ATCAGGTCTCCCTTTTCCCTCTGACATCTGACAGTACGACATCGGATTCACAGCTGTG 906  
 QY 221 HisAlaLysTyrGlnSerProTrrPheGlyTyrIleGluSerValGluValTyrThrTyr 240  
 Db CATGCTAAATACAGAGTCCCTGGTGGGAACCATGGAATCAGTTGAAGTCTACACTGG 966  
 QY 241 LeuAspAsnPheLeuLeuMetLeuGlyGlyIleProTrrGlnAlaTyrPheGlnArg 260  
 Db CTGTGAATATTTCTGTATTGATGCTGGGGGAATCCCATGGCAACCTTCCAGAGG 1026  
 QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGluValLeuSerPheLeuAlaIlePheGly 280  
 Db GTCTCTCTTCACTCCACGCACTTATGCTCAGGTACTGCTTCTCTGAGAGCTTTTGGG 1086  
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
 Db TGGCTGTGATAGCTCTACCCGCAATATGATGAGAGCTATTGTAGAGCTTCCAGAGCTGG 1146  
 QY 301 AsnGlnThrAlaTyrGlyTyrProAspPropheThyLysGluGluAlaAspMetIleLeu 320  
 Db AACCAAGCTGCTTACGGGATTCAGATCCCAAGACTTAAGAGGAGAGACATGATTTCTC 1206

QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
 Db CCGATGCTTCCAGTACAGTACCTGCTGCTGATCATCTCTTCTTGGGCTGTGCTGT 1266  
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360  
 Db TCAGCTGCTGATGCTCCACGCTGACTGCTCATCTCTCTGCGGAGTTCTATGTTGCT 1326  
 QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrrVal 380  
 Db CGGAATATCTACAGCTTCTTCTTCCAGCAAAATGATCAGACAGAAATTTGTGGGTC 1386  
 QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLys 400  
 Db ATGAGGATCATCGTGTGTTGTGTTCCGAGCATGCAACAGCCATGGCTTGTGACAGAG 1446  
 QY 401 ThrValTyrGlyLeuTrrPtyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
 Db ACTGTGTATGGGCTCTGATCCTGACCTGACCTGACCTGTCTACATCATCTTCCACAG 1506  
 QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
 Db CTGCTGTGTACTCTTCAACAGAACCAACACTTATGGGCACTTGTGTTATTT 1566  
 QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyProTyrLeuTyrLeuGlnProLeuIle 460  
 Db TTTGACTATCTCTGAGATTACTGGAGAGAGCCATCTATACCTTACCTTGGAGCCCTTATC 1626  
 QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPhePropheLys 480  
 Db TTTCTACCTCGTTATTAATCTGACAAAGATGATATTAACATCAGAGGTTCCCATTTTAA 1686  
 QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
 Db ACTCTCTCAAGTATACCTCATCTTACCAACATTTGTTTCTTATCTAGCCAGATAT 1746  
 QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520  
 Db CTATTTGAAAGGAAACCTTGGCTCCAAATTAAGATGATTTGATGCTGTGTCGAGAG 1806  
 QY 521 HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
 Db CACAGTGAAGAGAACTGAGCAAGACCATTCATGAGAAATGAAATTCAAATTAAT 1866  
 QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
 Db GAACTTGACCTGTGAAACCTGGCAGAGCTTACCTCAAGTTCACTTACCAATTAAG 1926  
 QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyTyrGluAspAsnLeuGln 580  
 Db GAGGCCCTCTCTGATTTATTTCACTCCGAGGGGCTTGGGACTGAAGTAACTTACAA 1986  
 Db 1927 GAGGCCCTCTCTGATTTATTTCACTCCGAGGGGCTTGGGACTGAAGTAACTTACAA 1986

RESULT 6  
 MM0401467 1743 bp mRNA linear ROD 16-AUG-2000  
 LOCUS Mus musculus mRNA for high affinity choline transporter (CHT1  
 DEFINITION gene).  
 ACCESSION AJ401467  
 VERSION AJ401467.1 GI:9843808  
 KEYWORDS CHT1 gene; high affinity choline transporter.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE  
 1 Wieland, A., Bonisch, H. and Bruns, M.  
 Molecular cloning of the human and murine high affinity choline  
 transporters and characterization of the human gene-structure  
 Unpublished  
 2 (bases 1 to 1743)  
 Bruns, M.  
 Direct Submission

JOURNAL Submitted (14-AUG-2000) Brnesh M., University of Bonn, Pharmacology  
and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
FEATURES Location/Qualifiers  
source 1..1743

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/cJ"  
/db\_xref="taxon:10090"  
/tissue\_type="Brainstem"  
1..1743  
/gene="CHT1"  
1..1743  
/gene="CHT1"  
/function="sodium- and chloride-dependent reuptake of  
chohline"  
/experiment="experimental evidence, no additional details  
recorded"  
/codon\_start=1  
/product="high affinity choline transporter"  
/protein\_id="CAC03719.1"  
/db\_xref="GI:9843809"  
/db\_xref="GOA:Q9BSW5"  
/translation="MSFVEGLVALILFYLFIIVGIVAAKTKNSGNPEHSEALIV  
GARDIGLVGFTMTATVVGXINGTAVAVYGGCGLAADAPYGLSLILGLF  
AKPMRSKVTMTLDPQOIYGRMGGLFIPALMGEMTRAAIFESALGHTISVITD  
VNISVVALAILLTVLGLYSAVTVVOLFCIFIGLWISVPALSHPAVDTIGFT  
AVHAKYQSPWLTIESVEVYTWLDFNLMLGLIPQVAYFORVLSSTAYAOVISYL  
AAGCLVVALPAICIGAGSDNMQNTAGYDPKTEADMLIPVLYQYLPVYSF  
FGLGVAASVMSADSSILASMSFARNIYOLSPRONSKDEIKVWMTITVVFASA  
TMAALLTKTVVGLWYLSLDLYIILPOLCVLPIKGNTVGAVAGYEGFLRTGG  
EPYILQPLIPYGYSDKNGINYNRPFKTISMTSPFTNICYSLAKYLFESSTLP  
FKLDVFDVAVARHSEENMDKITLVRENINKLIELAPVKEKROSLTUSFTPKKALLDV  
DSSPESGTEEDMLQ"

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 1743  
Score: 2958.00 Matches: 574  
Percent Similarity: 99.3% Conservative: 2  
Best Local Similarity: 99.0% Mismatches: 4  
Query Match: 98.8% Indels: 0  
DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x MMU401467 (1-1743)

QY 1 MetProPheHISValGluGlyLeuValAlaIleIleLeuPheTyLeuLeuIlePheLeu 20  
DB 1 ATGCTTTCCAGTGAAGAGACTGGTAGCTATTATCTCTTACCTCTTATATTTCTG 60  
QY 21 ValGlyIleTPAlaAlaIleTPlysThrIysAsnSerGlyAsnProGluIUAArgSerGlu 40  
DB 61 GTTGAATATGGGCTGCATGAAAAACAAAACAGGGGCAACCCAGAAAGACAGATGAA 120  
QY 41 AlaIleIleValGlyIYArgAspIleGlyLeuLeuValGlyIYPhethrMetThrAla 60  
DB 121 GCCATCATAGTGGGGGCGGTGACATGGTTGTTGGTGGTTTACCATGACAGCC 180  
QY 61 ThrTPValGlyIYTyriIleasnGlyThrAlaGluAlaValTyriGlyProGlyCys 80  
DB 181 ACCTGGTTGGAGAGGCTACATCAATGAGCAGCAGTAGAGTATGGCCAGGTTT 240  
QY 81 GlyLeuAlaIleTPAlaHisAlaProIleGlyTyriSerIleuSerIleuIleleuGlyIleu 100  
DB 241 GGTCTAGCTGGGCTCAGGCAACCATTTGATATCTCTAGCTTAATTTTAACTGGTCTG 300  
QY 101 PhePheAlaIysProMetArgSerIysGlyTyriValThrMetIleuAspProPheIysGln 120  
DB 301 TTTTTCGGAAACCTATGCGTTCCAAAGGATATGACATATGTTACACCATTTCAACG 360  
QY 121 IleTyriGlyIysArgMetGlyIYLeuLeuPheIleProAlaLeuMetGlyIYUetPhe 140  
DB 361 ATCTATGAAAGCGATGGGTGGCTGCTTCACTCCGCACTGATGGGAGAGATGTTCC 420

QY 141 TPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
DB 421 TGGGCTGAGCAATTTTCTGTCATTTAGGGGCCACCATCAGCGATCATGATGATGATG 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyriThrLeuValGlyIY 180  
DB 481 GTGAACATATCGGTATATGTCCTCGCACTCATGCTATTTTATACCTTATGAGTGG 540  
QY 181 LeuTyriSerValAlaTyriThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200  
DB 541 CTCTACTGTGGAGATATATGATGATGTTCCAGCATTTCTGATTTTATAGACTGTG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
DB 601 ATCACTGTCCTTTTGGCCCTGTCATCCTCAGACACGACCATCGATTCACACCTTG 660  
QY 221 HisAlaIysTyriGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyriThrTrp 240  
DB 661 CATGCTTAATATCCAGAGTCCCTGGCTGGGAACCATTTGAATCATGTTGAAGTCAACCTG 720  
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIYIleProTrpGlnAlaTyriPheGlnArg 260  
DB 721 CTGTATATTTTCTGTATATGATGCTGGGTGAATCCCATGGCAAGCCTTACAGAGG 780  
QY 261 ValLeuSerSerSerSerAlaThrTyriAlaGlnValLeuSerPheLeuAlaAlaPheGly 280  
DB 781 GTCCCTCTTATCTCTCAGCAACCATTCCTCAGGATACGTCTCTTGGCAGCTTTTGG 840  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300  
DB 841 TGCCGTGATAGGCTCTACCCGCCATATGATAGAGATATTTGAGACTTCCACAGACTGG 900  
QY 301 AsnGlnThrAlaTyriGlyTyriProAspProIysThrIysGluGluAlaAspMetIleLeu 320  
DB 901 AACCAAGCTGCTACGAGGTATCCAGATCCAAAGCTAAGAGAAAGACAGATGATTC 960  
QY 321 ProIleValLeuGlnTyriLeuCyProValTyriIleSerPhePheGlyLeuGlyAlaVal 340  
DB 961 CCGATGCTTCCAGATATCTTCCCTGCTGTGACATCTCTTCTTGGGCTGGTGCTGT 1020  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
DB 1021 TCAGCTGCTGTCATGTCCTCAGCTGACTGTCACATCCCTGCGGCAAGTTCTATGTTGCT 1080  
QY 361 ArgAsnIleTyriGlnLeuSerPheArgGlnAsnAlaSerAspIysGluIleValTrpVal 380  
DB 1081 CGGAATATCTACCAAGCTTCTCTTCAGACAAATGATCAGACAAAGAAATGTGTGGGCTC 1140  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIys 400  
DB 1141 ATGAGGATCACGTGCTGTGTGGAGCATCTGCAAGCCATGCGCTTGTGTGACGAAG 1200  
QY 401 ThrValTyriGlyLeuTrpTyriLeuSerSerAspLeuValTyriIleIleIlePheProGln 420  
DB 1201 ACTGCTATGGGCTGTGATCCTGACCTGACCTGTGCTCATCATCATCTTCCACAG 1260  
QY 421 LeuLeuCyValLeuPheIleIleIysGlyTyriAsnThrTyriGlyAlaValAlaGlyTyriIle 440  
DB 1261 CTGCTCTGTGACTCTTTCATCAAGAAACCAACTTATGGGGCAGTGTGCTGTTATATT 1320  
QY 441 PheGlyLeuPheLeuArgIleThrGlyIYIleGluProTyriLeuTyriLeuGlnProleuIle 460  
DB 1321 TTGGACTATTCCTGAGATTTATCTGAGAGAGCCATATCTATATTGTCAGCCCTTATTC 1380  
QY 461 PheTyriProGlyTyriTyriSerAspIysAsnGlyIleTyriAsnGlnArgPheProPheIys 480  
DB 1381 TTCTACCTGCTGTTATATCTGACAGAAATGTTATATCATTCAGAGGTCCCATTTTAA 1440  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyriLeuAlaIysTyri 500  
DB 1441 ACTCTCTCAATGTTACCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAAAGTAT 1500

QY 501 LeuPheGluSerGlyThirleuProProlySleuAspValPheAspAlaValAlaArg 520  
 Db 1501 CTAATTGAAGAGGAGACCTTGCTCCAAATTAATGATGATTTGATCTGTGTGCGAAG 1560  
 QY 521 HisSerGluGluAsnMetAspLysThrIleValAlaArgAsnGluAsnIleLysLeuAsn 540  
 Db 1561 CACAGTGAAGAGAACATGACACAGACCAATCTAGTCAGAAATGAAATATCAAAATTTAAAT 1620  
 QY 541 GuleuAlaProValLysProArgGlnSerLeuThirleuSerSerThrPheThrAsnLys 560  
 Db 1621 GAACCTTGACCTGTGAACCTCGGACAGCTTAACCTCACTTCACTTCAACCAATAG 1680  
 QY 561 GAlaAlaLeuLeuAspValAspSerSerProGluGlySerGlyThirGluAsnLeuGln 580  
 Db 1681 GAGGCCCTCTTGATTTGATTTCACTCCGAGGGGCTGGAGCTGAAGATTAATTACAA 1740  
 RESULT 7  
 BD012718 1743 bp DNA linear PART 02-AUG-2002  
 LOCUS High-affinity choline transporter.  
 DEFINITION BD012718  
 ACCESSION BD012718.1 GI:22092907  
 VERSION WO 0116315-A/2.  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Rattus.  
 1 (bases 1 to 1743)  
 HAGA,T. and Okuda,T.  
 High-affinity choline transporter  
 Patent: WO 0116315-A 2 08-MAR-2001;  
 JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA  
 COMMENT OS Rattus norvegicus (rat)  
 PN WO 0116315-A/2  
 PD 08-MAR-2001  
 PR 18-AUG-2000 WO 2000JP005545  
 PF 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI  
 TATSUYA HAGA, TAKASHI OKUDA  
 PC C12N5/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC  
 A61K38/17,  
 PC A61K45/00, A61P25/28, G01N33/53, A01K67/027  
 CC  
 FH Key Location/Qualifiers  
 FT CDS (1). (1743).  
 Location/Qualifiers  
 1. 1743  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 1743  
 Score: 2956.00 Matches: 572  
 Percent Similarity: 99.3% Conservative: 4  
 Best Local Similarity: 98.6% Mismatches: 0  
 Query Match: 98.8% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-724-806-4 (1-580) x BD012718 (1-1743)  
 QY 1 MetProPheHisValGlyLeuValAlaIleIleLeuPheThirleuLeuIlePheLeu 20  
 Db 1 ATGCTCTTCATTTAGAGACTAGTAGCATTAATCTGTTCACTCTTTATATTTCTG 60  
 QY 21 ValGlyIleThrAlaIleTrpLysTrpLysAsnSerGlyAsnProGluGluArgSerGlu 40  
 Db 61 GTTGGATATATGGGCTGCATGAGAAACCAAAACAGCGGTAAATGACAGAGAACCGAGAA 120  
 QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
 Db 121 GCCATATAGTTGGGGGCCGAGACATTGTTGTGTGGTGGTTTATACATGACAGCC 180

QY 61 ThrTrpValGlyGlyGlyThrIleAsnGlyThrAlaGluAlaValTrpGlyProGlyCys 80  
 Db 181 ACTGGCTTGGAGAGAGTTACATCAACGGGACAGCTGAAGCATTTATGGGCCAGGTTGT 240  
 QY 81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTrpSerLeuSerLeuIleLeuGlyLeu 100  
 Db 241 GGTCTAGCTTGGGCTCAGGCACCCATGGATTTCTGAGCTGATTTTATAGCTGGCCCTG 300  
 QY 101 PhePheAlaLysProMetArgSerLysGlyTrpValThirMetLeuAspProPheLysGln 120  
 Db 301 TTTTGTGCAAAACCTATGGGCTTCCAGGAGATATGACTATGTATGACCCGTTTCAACG 360  
 QY 121 IleTrpGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyLysMetPhe 140  
 Db 361 ATCTATGAGAAAGCGATGGGTGGGCTGCTTCACTCCGCACTGATGGAGAGATGTTT 420  
 QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db 421 TGGGCTGACAGCAATTTTCTCTGATTAAGGGGCTACCATCAGCTAATCATTTGATGTGAT 480  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrpThirleuValGly 180  
 Db 481 GGTGAACATACGGTCAATGTCTCCGACCTCATTTGCAATCTTATACCTCGTGGAGAGG 540  
 QY 181 LeuTrpSerValAlaTrpThirAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200  
 Db 541 CTCTACTCTGTGGCATATATCTGATGTGTACAGCTATTCGCACTTTTATAGATTTCTGG 600  
 QY 201 IleSerValProPheAlaLeuSerHisProAlaValThirAspIleGlyPheThrAlaVal 220  
 Db 601 ATCAGAGTCCCATTTGCTCCGTGACATCTGCAGTACGACGACATGTGATTCACCTGCTG 660  
 QY 221 HisAlaLysTrpGlnSerProTrpLeuGlyThrIleGluSerValGluValTrpThrTrp 240  
 Db 661 CATGCTAAATACAGAGTCCCTGGCGGAGAACATTGAATCACTTGAAGCTTACACCTGG 720  
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTrpPheGlnArg 260  
 Db 721 CTTGATATTTTCTGTGTGTGATGCTGGGTGGAATACATGGAGACCTTCCAGAGG 780  
 QY 261 ValLeuSerSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
 Db 781 GTCCTCTCTTCATCGTCAGGACCTATGCTCAGGTGCTGCTCTCGGACGCTTTGGG 840  
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrTrp 300  
 Db 841 TGCCTGTGATGCTCTTACAGCCATTTGCATTTGGGGCCATTGGACCTTCCAGAGCTGG 900  
 QY 301 AsnGlnThrAlaTrpGlyTrpProAspProLysThrLysGluGluAlaAspMetIleLeu 320  
 Db 901 AACCAATCTGCATATGGGTTTCCAGATCCCAAGACCAAGAGAGACATGATTTCTC 960  
 QY 321 ProIleValLeuGlnTrpLeuCysProValTrpIleSerPhePheGlyLeuGlyAlaVal 340  
 Db 961 CCGATTTGTTCTACAGTACCTGCTCGTGTGATATTTCTTGTGGGCTGTGCTGTT 1020  
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360  
 Db 1021 TCTGCTGCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 QY 361 ArgAsnIleTrpGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380  
 Db 1081 CGGAATATCTACAGCTTCTTCTTCAACAATAAGCATGACAGCAAGAAATGTGTGGTCT 1140  
 QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
 Db 1141 ATGAGGATCATCTGT 1200  
 QY 401 ThrValTrpGlyLeuTrpTrpTrpLeuSerSerAspLeuValTrpIleIlePheProGln 420  
 Db 1201 ACTGTATATGGGCTCTGTGATCTGAGCTTGAACCTTGTCTTACATCATCTTCCACAG 1260

QY	421	LeuenuCysValIleuPheHlleYsGlyYThrSerThrYrrGylAvalAlaGlyTyrTlle	440
Db	1261	CTGCTCTGTGACTCTTATCATAAAGAAACCAACTTATGGGGAGATTGCTGGTTAATT	1320
QY	441	PheGlyLeuPheLeuArgIleThrGlyGlyGluProYrYLeuYrYLeuGlnProLeuIle	460
Db	1321	TTTGGAAGTTTTCCTGAGAAATTACCGGAGGAGGCATATCTATATCTTGACGCTTATC	1380
QY	461	PheYrYrProGlyYrYrYrSerAspYsAsnGlyIleYrAsnGlnAnrGpheProPheYls	480
Db	1381	TTCTAACCTCGGTATTTACCTTGACACAGATGATATATCAATCAAGAGGTTCCATTAAA	1440
QY	481	ThrLeuSerMetValThrSerPheThrAsnIleCysValSerYrYLeuAlaYsTyr	500
Db	1441	ACTGCTCTCAAGTTTAACTCATTTCTTTACCAACATTGTGTCTCTATCTAGCCAAAT	1500
QY	501	LeuPheGluSerGlyYThrLeuProProYlsLeuAspValPheAspAlaValAlaArg	520
Db	1501	CTATTGTGAAGTGGAAACCTTGCTCCAAAATTAGATATATTTGATGGCTGTCTCAAG	1560
QY	521	HisSerGluGluAsnMetAspYsThrIleLeuValArgAsnGluAsnIleYsLeuAsn	540
Db	1561	CACAGTGAAGAGACATGACAGACCCATTCTAGTCAGAAATGAAATACATATAAT	1620
QY	541	GluLeuAlaProValYsProArgGlnSerLeuThrLeuSerSerThrPheThrAsnYls	560
Db	1621	GAAGTTCACCTGTAAACCTCTGACAGAGCCTAACCTCAGTTCACTTCCACCAATAA	1680
QY	561	GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyYThrGluAspAsnLeuGln	580
Db	1681	GAGGCTCTCTTGATGTGATTCCAGTCCAGAGGAGATCTGGAGACTGAAGATTAACCTCA	1740
RESULT 8			
LOCUS	E49870	1743 bp	DNA
DEFINITION	High-affinity choline transporter.	linear	PAT 27-AUG-2002
ACCESSION	E49870		
VERSION	E49870.1	GI:22554901	
KEYWORDS	JP 2001136976-A/2.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 1743)		
TITLE	Haga, T. and Okuda, T.		
JOURNAL	High-affinity choline transporter Patent: JP 200136976-A 2 22-MAY-2001; SCIENCE & TECH AGENCY		
COMMENT	OS Rattus sp. (rat) PN JP 200136976-A/2		
	PD 22-MAY-2001		
	PF 27-DEC-1999 JP 1999368991		
	PI TATSUYA HAGA, TAKASHI OKUDA		
	PC C12N15/09, A01K67/027, A61K38/00, C07K14/47, C07K16/18, C07K19/00,		
	PC C12N5/10,		
	PC C12P21/02, C12P21/08, C12Q1/00, C12N15/00, A61K37/02, C12N5/00 CC		
	FT Key		
	Location/Qualifiers		
	(1)..(1743).		
FEATURES			
source	1..1743		
	/organism="Rattus sp."		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10118"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	1743
Score:	2956.00	Matches:	572
Percent Similarity:	99.3%	Conservative:	4
Best Local Similarity:	98.6%	Mismatches:	4
Query Match:	98.8%	Indels:	0
DB:	2	Gaps:	0

	US-10-724-806-4 (1-580) x E49870 (1-1743)	
QY	1 MetProhEhiSaValgiuglyLeuValAlaIleIleuBheTyriLeuLeuIlePheLeu	20
Db	1 ATGCTTTCATGATAGAAGACGTAGTGGCATATATCTGTTCTTACCTCTTATATTTCTG	60
QY	21 ValGlyIleTrrPaAlaIaTrrPlySThrLyAsnSerGlyAsnProGluGlyuAspSerGlu	40
Db	61 GTTGAAATATGGCTGTGATGAGAAACCAAAACGCGGATATGCAGAGAAACGAGGAA	120
QY	41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThraAla	60
Db	121 GCCATCATAGTGGGGGCCGAGACATTTGGTTGGTTGGTTTGTGGTTTACATGACAGCC	180
QY	61 ThrTrpValGlyGlyGlyTyriIleAsnGlyThrAlaGluAlaValTyriGlyProGlyCys	80
Db	181 ACCTGGGTGGAGAGGTTACATCAACGGGACAGCTGAACAGATTATGGGCCAGTTGT	240
QY	81 GlyLeuAlaTrrPaAlaIleAsnProIleGlyTyriSerLeuSerLeuIleLeuGlyGlyLeu	100
Db	241 GGCTTACTTGGGGCTCAGGACCCATTTGAAATCTCTGAGTCGATTTAAAGGGGCTG	300
QY	101 PhePheAlaIlyPsrPromeCtArgSerLySGlyTyriValThrMetLeuAspPropheLySgl	120
Db	301 TTTTTCGAAACCTATAGCGCTTCCAAAGGATATGTGACTATGTTAAGCCCGTTTCAACAG	360
QY	121 IleTyriGlyIlySarMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe	140
Db	361 ATCTATGGAACCCCATAGTGGGGCTGCTGTCTTCCTCCACATGAGGAGAGATGTTTC	420
QY	141 ThrPaAlaAlaIlePheSerAlaLeuGlyValaThrIleSerValIleIleAspValAsp	160
Db	421 TGGCTGTGAGAAATTTCTCTGCAATGGGGCTTACCACTACGCTTAATCATTTGATGGAT	480
QY	161 ValAsnIleSerValIleValIleSerAlaLeuIleAlaIleLeuTyriThrLeuValGlyGly	180
Db	481 GTGAACATATGGTCATTTGCTCCGCACTCATTTGCCATTTCTTATACCTCGTGGAGGG	540
QY	181 LeuTyriSerValaIaTyriThrAspValaGlnLeuPheCysIlePheIleGlyLeuTrp	200
Db	541 CTCTACCTCTGGGCAATACGATGAGTGTGACAGCTATTTCTGCAATTTTAAAGATGTGG	600
QY	201 IleSerValProPheAlaLeuSerHisProAlaValaThrAspIleGlyPheThrAlaVala	220
Db	601 ATCAGTGTCCATTTGCCCTGTCACTCCGCACTCCACATTTGATTCACCTGCTGTG	660
QY	221 HisAlaIlyTyriGlnSerProTrrPheGlyThrIleGlnSerValGluValTyriThrTrp	240
Db	661 CATGCTAAATACCAAGAGTCCCTGGCTGGAAACATTGATCAGTTGAAGTCACTACCTGG	720
QY	241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrGlnaIaTyriPheGlnArg	260
Db	721 CTTCATATATTTTCTGTTGTGATGCTCTGGTGGATACCATGCGAAGCTCTTCCAAAGG	780
QY	261 ValLeuSerSerSerSerAlaThrTyriAlaGlnValLeuSerPheLeuAlaIlePheGly	280
Db	781 GTCTCTCTTTCATGCTCAGCGAAGCTTAAGCTCAGGTGCTGCTCTTCTTGGGAGCTTTGGG	840
QY	281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp	300
Db	841 TGCTGTGGATGGCTTACCAAGCATTTTGCATTGGGGCCATTGGAGCTCCACAGACTGG	900
QY	301 AsnGlnThrAlaTyriGlyTyriTrpAspPropolysThrIlySglGluAlaAspMetIleLeu	320
Db	901 AACCAAACTGATATGGGTTTCCAGATCCCAAGACCAAGAGGAAAGACATGATTTCTC	960
QY	321 ProIleValLeuGlnTyriLeuCysProValTyriIleSerPhePheGlyLeuGlyAlaVala	340
Db	961 CCGATTGTTCTACAGTACCTCTGCCCTGTGTACATTTCTTCTTGGGCTTGGTGTGCTT	1020
QY	341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla	360

Db	1021	CTGTGCTGTGCATGATCTCTCGGCTGACATCATCATCATCATGACAGAACTTCATGTTTGGCT	1089
Qy	361	ArgSnnlleYrGlnLeuSerPheArgGlnAsnAlaSerAspLyGluIleValITrpVal	380
Db	1081	CGGAATATCTCAACACCTTTCCTTCAGCAAAATGATCAGCAAGAAATTGTGGGCTC	1140
Qy	381	MetArgIleTrpValIleuValPheGlyAlaSerAlaTrpAlaMetAlaLeuLeuThrIys	400
Db	1141	ATGAGATCATCTGTGTGTGTGTGGACATCTGCACAGCCATGCGCTTGTCTACGAG	1200
Qy	401	ThrValITyGlyLeuTrpITyPheLeuSerSerAspLeuValITyIleIleIlePheProGln	420
Db	1201	ACTGTGTATGGGCTCTCGGTACCTGAGCTGACCTGACCTGTCTCATCATCATCTTCCACAG	1260
Qy	421	LeuLeuCySValIleuPheIleIleGlyIThrSerTrpITyGlyValAlaValAlaGlyITyTrle	440
Db	1261	CTGCTCTGTGTACTTTCATCAAAAGAAACCAACACTTAATGGGCGAGTGTGGTATATTT	1320
Qy	441	PheGlyLeuPheLeuArgIleThrGlyGlyGluProITyLeuITyLeuGlnProIleuIle	460
Db	1321	TTTGGACTTTTTCCTGAGATATTACGGAGAGAGCGCATTCATCTCATCTTCACCCCTTATTC	1380
Qy	461	PheITyProGlyITyITyITySerAspIysAsnGlyIleIyIysAsnGlnArgPheProPheIys	480
Db	1381	TTCTAACCTCGTATTATACCTCGACAGAAATGATATCATCATCAGAGCTTCCCAATTTTAA	1440
Qy	481	ThreuSerMetValIThrSerPhePheThrAsnIleCySValSerITyLeuAlaIlySTyR	500
Db	1441	ACTCTCTCCATGGTTACTCTTATCTTTACCAACATTGTGTGTTCTTACTTACGCAAGTAT	1500
Qy	501	LeuPheGluSerGlyIThrLeuProProIysLeuAspValPheAspAlaValAlaAlaArg	520
Db	1501	CTATTGTGAAGTGAACCTTGCTGCCTCAAAATTAGATATATTGATGCTGTGTCTCAAG	1560
Qy	521	HisSerGluGluAsnMetAspIysThrIleIleuValAlaGAsnGluAsnIleIysLeuAsn	540
Db	1561	CACAGTGAAGGAAACATGACAGAACGCCATTTGATGACAGAAATGAAMACATCAAAATTAAAT	1620
Qy	541	GluLeuAlaProValIysProArgGlnSerLeuIThrLeuSerSerIThrPheThrAsnIys	560
Db	1621	GAACTGTGACCTGTAAACCTCGACAGAGCCTTAACCTCACTTACCTTACCAATATAA	1680
Qy	561	GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyIThrGluAspAsnLeuGln	580
Db	1681	GAGGCTCTCCTTGATTGTGATTCCAGTCCAGAGGGATCTGGAGACTGAAGATTAACCTTACA	1740
RESULT 9			
AB030947			
LOCUS	AB030947	4904 bp	mRNA linear ROD 03-FEB-2000
DEFINITION	Rattus norvegicus mRNA for high-affinity choline transporter CHT1, complete cds.		
ACCESSION	AB030947		
VERSION	AB030947.1	GI:6863033	
KEYWORDS	choline transporter; high-affinity choline transporter CHT1.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Rattus.		
AUTHORS	Okuda,T., Haga,T., Kanai,Y., Endou,H., Ishihara,T. and Katsura,I.		
TITLE	Identification and characterization of the high-affinity choline transporter		
JOURNAL	Nat. Neurosci. 3 (2), 120-125 (2000)		
PUBMED	10649566		
REFERENCE	2 (bases 1 to 4904)		
AUTHORS	Okuda,T.		
TITLE	Direct Submision		
JOURNAL	Submitted (09-AUG-1999) Takashi Okuda, University of Tokyo, Faculty of Medicine, Department of Neurochemistry, Hongo 7-3-1, Bunkyo-ku 113-0033, Japan (E-mail:okudaem.u-tokyo.ac.jp, Tel:+81-3-5841-3560, Fax:+81-3-3814-8154)		
COMMENT	Sequence updated (11-Jan-2000).		

FEATURES

source

location/Qualifiers

1..4904

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Mistar"

/db\_xref="taxon:10116"

/clone="CHT1"

/issue\_type="spinal cord"

/clone\_lib="rat spinal cord cDNA library"

/dev\_stage="adult"

224..1966

/codon\_start=1

/product="high-affinity choline transporter CHT1"

/protein\_id="BAA90484.1"

/db\_xref="GI:6863034"

/translation="MPHHVGLVAIIIFLLIFLVGIVGAAWAKTKSGNMEEREBAIIVGGRIIDLVGGFTWTATWVGAGYTINGAEVAVDPGGLAAQAPIGISUSLIGLGFAPKPRSKYVATMLDPPQOIKGRMGGLFIPALMGEMFMAALFSAIGATI SVIIVDVUNISVISAVALIYTLVGLGVSAVADIVDLFCIFGLWISVPFALSHPAVDIGFTAAHAKQSPWIGTIESVEVYTMIDNFIILMLGGLIPQWYFORVLSSSATAYAOVLSFPAARCLVMAIPALICIGATIGASTDMNOCVAFEPDPKTEAEADMLTIVLOYLCVYISFPGCAVSAAMVSSADSIISSASMEANIVOLS.FRONASDKELVIMVRILFVPRGSA.TMAALTKTYIVGLMYLSSDLYITIIIFQLLCVLPKIKNTYGAAGTICELPLRITGAEPLVYLOPLIFYPGYPDPKNGIYNQRPFTLSMVS.FFNICVSYIARYL.FESGLPKDIDFPAVVSRRSEENMDKTIILVRNENIKLNLAPVYKPROSLTSLSTFNKALDLVDSSPEGSGETEDNIQ"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4904

Score: 2956.00 Matches: 572

Percent Similarity: 99.3% Conservative: 4

Best Local Similarity: 98.6% Mismatches: 4

Query Match: 98.8% Indels: 0

DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x AB030947 (1-4904)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleuPheTyrlLeuLeuIlePheLeu 20

Db 224 ATGCTTTCCATGATGAGAAGGACTAGTAGCGATTATCCTGTTCACTCTTATATTTCTG 283

QY 21 ValGlyIleTrpAlaIleatTrpLysThrLysAsnSerGlyAsnProGluGluAsnSerGlu 40

Db 284 GTTGAAATATGGCTGCATGAGAAAACCAAAACACGCGTATGACAGAAAGACGCGAA 343

QY 41 AlaIleIleValGlyGlyIleArgAspIleGlyLeuLeuValGlyGlyPheTrpMetThAla 60

Db 344 GCCATCATAGTTGGGGGCCCGAGACATTGGTTTGTGGTGGTTTACCATGACAGCC 403

QY 61 ThrTrpValGlyGlyGlyTyrlIleAsnGlyThrIleAspIleValIleValTyrglyProGlyCys 80

Db 404 ACTGGGGTGGAGGAGGTTACATCAACGGGACAGCTGAACAGATTATGGGCCAGGTTGT 463

QY 81 GlyIleuIleatTrpAlaHisAlaProIleGlyTyrSerIleuSerIleuIleLeuGlyGlyLeu 100

Db 464 GGTTACCTGGGGCTGACGGAACCCATTGGATATCTTGTAGTTCGATTATGAGTGGGCTTG 523

QY 101 PhePheAlaLysProMetArgSerLeuGlyTyrValThrMetLeuAspProPheLysGln 120

Db 524 TTTTTCGAAAACCTTATGCGGTTCCAAAGGATATGTGACTATGTTAGACCCGTTTCAACAG 583

QY 121 IleTyrglyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140

Db 584 ATCTATGAGAAACCCATGAGGGCTGCTGTTCACTCCCTCAGCATGAGGAGAGATGTC 643

QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160

Db 644 TGGGCTCAGCAATTTTCTCTGCAATTAGGGGCTACCATCAGCGTAATCATTTGATGTGAT 703

QY 161 ValAsnIleSerValIleValIleSerAlaLeuIleAlaIleLeuTyrrThrLeuValGlyGly 180

Db 704 GTGAACATATGGCATATGTTCTCCGACCTAATGGCAATCTTTATATACCTCGTGGAGGGG 763

QY 181 LeuTyrSerValAlaIleTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200  
Db CTCTACTCTGGGCGATATCTGATCTTGTAACAGTATTCTGATTTTATATAGATTGG 823  
QY 201 ILeSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220  
Db ATCACTGCCATTTGGCCCTGTCACATCTGCAGTCACGACCATTTGGATTCTGCTGG 883  
QY 221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGluSerValGluValTyrThrTrp 240  
Db CATGCTAATATACCAAGTCCCTGGGAGCAACATTGAATCATGTTGAGTCTACACCTGG 943  
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260  
Db CTGTATATATTTCTGTGTGTGATGCTGGGTGGAAATACCATGCAAGCCACTTCCAGAG 1003  
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
Db GTCTCTCTCTCAATGCTCAGCGACCTATGCTCAGGTGCTCTCTCTGCGACCTTTGG 1063  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300  
Db TGCCGTGGATGGCTCTACCAACCATTTGSCATTGGGGCCATTGGAGCCTCCACAGACTGG 1123  
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320  
Db AACCAACCTGCATATGGGTTCACATCCCAAGACCAAGAGAGCAACATGATTC 1183  
QY 321 ProIleValIleGlnTyrLeuGlnCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
Db CGGATTTGTTCTACAGTACCTGCTGCTGTCATTTCTTCTTGGGCTGGTGTGTT 1243  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
Db TCTGCTGCTGATGATGCTCGGCTGATCATCATCTATACAGCAAGTTCCATGTTTGT 1303  
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal 380  
Db CGGAATATCTACCACTTCCTTCACAGCAAAATGATCATCAGACAGAAATGCTGGGTG 1363  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
Db ATGAGATATCACTGTTGTTGTTGTTGAGCATCTGCAACGACCTGCTGCTCAAGAG 1423  
QY 401 ThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
Db ACTGATGATGGGCTCTGGTACCTGAGCTGACCTTGTCTACATCATCATCTTCCACAG 1483  
QY 421 LeuLeuGlyValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
Db CTGCTCTGTGATCTTCAATCAAAAGAACCACTTATGGGGCAGTGTCTGTATATAT 1543  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460  
Db TTTGACCTTTTCTAGAAATTCACGAGAGAGCCATATCTATCTTGCAGCCCTTAATC 1603  
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPhePheProLys 480  
Db TTTCTACCTGGTTATTAACCTGACAAATGATATATACATCAGAGGTTCCCATTTAA 1663  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500  
Db ACTCTCTCATAGTATACCTCATCTTTCACCACTTTGTTGTTCTATCTTACCAAGAT 1723  
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520  
Db CTATTGAAAGTGAACCTTGCCTCCAAATATAGATATATTTGATGCTGTTGTTCAAG 1783  
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
Db CACAGTGAAGAGAACATGAGACAAACATTTCTAGTACAGAAATGAAAAACATCAATTAAT 1843

QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
Db GAACCTTGACACCTGTAAGCCCTCGACAGAGCCTAACCTTCATCACTTACCAATATA 1903  
QY 561 GluAlaLeuLeuAspValAspSerSerProGluLysGlyTyrGlnLysAsnLeuGln 580  
Db GAGGCTCTCTGATGTTGATTCAGTCCAGAGGAGATCTGGGACTGAAGATTAATTACA 1963

RESULT 10  
BD012719 1743 bp DNA linear PAT 02-AUG-2002  
LOCUS BD012719  
DEFINITION High-affinity choline transporter.  
ACCESSION BD012719.1 GI:22092908  
VERSION BD012719.1  
KEYWORDS WO 0116315-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1743)  
AUTHORS Haga, T. and Okuda, T.  
TITLE High-affinity choline transporter  
JOURNAL Patent: WO 0116315-A 3 08-MAR-2001;  
JAPAN SCIENCE AND TECHNOLOGY CORP, TATSUYA HAGA, TAKASHI OKUDA  
COMMENT OS Homo sapiens (human)  
PN WO 0116315-A/3  
PD 08-MAR-2001  
PF 18-AUG-2000 WO 2000JP005545  
PR 27-AUG-1999 JP 99P 240642, 27-DEC-1999 JP 99P 368991 PI  
TATSUYA HAGA, TAKASHI OKUDA  
PC C12N15/12, C07K14/47, C12Q1/68, C07K19/00, C07K16/18, C12N5/10, PC  
A61K38/17,  
PC A61K45/00, A61P25/28, G01N33/53, A01K67/027  
CC  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1..1743  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 1743  
Score: 2795.00 Matches: 537  
Percent Similarity: 96.64 Conservative: 23  
Best Local Similarity: 92.64 Mismatches: 20  
Query Match: 93.48 Indels: 0  
DB: 2 Gaps: 0

US-10-724-806-4 (1-580) x BD012719 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
Db ATGCTTTCATGTGAAGGACTGATAGCTATCATCGTCTTCACTTAAATTTGCTG 60  
QY 21 ValGlyIleTyrPalaIleTyrLysAsnSerGlyAsnProGluGlnArgSerGlu 40  
Db GTTGGAAATATGGGCTGCTGAGAACCAAAACAGTGCACGCGAGAGACCGCGAA 120  
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
Db GCCATCATAGTTGGTGGCCGAGATATGTTATGTTGGTGGATTACCATGACACCT 180  
QY 61 ThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGlnAlaValTyrGlyProGlyLys 80  
Db ACTGGGTGCGAGAGGGGTATATCATGTCACAGCTGAAGAGATTATGATCCAGTTAT 240  
QY 81 GlyLeuAlaTyrPalaIleProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
Db GGCCTAGCTTGGGCTCAGGACCAATTGGATATCTCTTACTGTGATTTTGGTGGCTG 300

QY 101 PhePheAlaIySerProMetArgSerIyTyValThrMetLeuAspProPheIySgIn 120  
 Db 301 TTCTTGGCAAAACCTTATGCGTTCAAAGGGGTATGTACATGTTAAGCCGTTTCAGCAA 360  
 QY 121 ILeTyGlyLysArgMetGlyIyLeuPheIleProAlaLeuMetGlyIyMetPhe 140  
 Db 361 ACTATGAGAAAGCATGGGGCGGACCTCGTTTATTCGTGACACTGATGGAGAAATGTTTC 420  
 QY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db 421 TGGGGCGAGCAATTTCTCTGCTTGGAGACCAATCAGCGTGTATCATCGATGGAGAT 480  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyThrIleValIyGly 180  
 Db 481 ATGCACATTTCTGTCTCATCTCTGCACTCATTCGACACTGTGACACACTGGTGGAGGG 540  
 QY 181 LeuTySerValAlaTyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
 Db 541 CTCTATTCTGGGCTTACACTGATGTGCTTACGCTCTTTTGGCATTTTGTAGGGCTGTGG 600  
 QY 201 ILeSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 Db 601 ATCAGGCTCCCTTGGCATTTGTCAATCCTGACGTGCAGACATCGGGTTCACCTGCTGTG 660  
 QY 221 HisAlaIyTyGlnSerProTyrLeuGlyIyThrIleGlnSerValGlnValTyThrTyr 240  
 Db 661 CATGCCAAATACCAAAAGCGGTGGCTGGAGACTGTGACTCATCTGAAGCTTACCTTGG 720  
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIyIleProTyrGlnAlaTyThrPheGlnArg 260  
 Db 721 CTGTAGATGTTTCTGTGTTGATGCTGGGGAGATCCCATGGACATGACATCTTTCAGAG 780  
 QY 261 ValLeuSerSerSerSerAlaThrTyThrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
 Db 781 GTTCTCTCTCTCTCCACGCCACATTCATGCTGCTGCTCTCTCTCTCTCTCTCTCTCT 840  
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
 Db 841 TGGCTGTGATGGCCATCCAGGCATATCTATGGGGCCATTTGGAGCATCAACAGCTGG 900  
 QY 301 AsnGlnThrAlaTyGlyTyProAspProIySgIyThrLysGlnGlnAlaAspMetIleLeu 320  
 Db 901 AACCAAGATGCATATGAGGCTTCCAGATCCCAAGATCAAGAAAGGACAGCATGATTTTA 960  
 QY 321 ProIleValLeuGlnTyLeuCysProValTyThrIleSerPhePheGlyLeuGlyAlaVal 340  
 Db 961 CCATATGTTCTGCAGATATCTGCTGCCCTGTATATATTTCTTTCTTGTGTGTGACAT 1020  
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360  
 Db 1021 TCTGTGCTGTATATGATCATGACATCTTCCATCTGTGACAAAGTTCCATGTTTGA 1080  
 QY 361 ArgAsnIleTyGlnLeuSerPheArgGlnAsnAlaSerAspIySgIyLeuValTyThrVal 380  
 Db 1081 CGGAACATCATACAGCTTCTTCCCTTCAAGAAATGCTTCGGAACAAGAAATCGTTGGGTT 1140  
 QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuTyThr 400  
 Db 1141 ATGCCAATCAGATGTTTGTGTTTGGAGATCTGCACAAAGCCATGCGCTTGCAGAGAAA 1200  
 QY 401 ThrValTyGlyLeuTyThrIleuSerSerAspLeuValTyThrIleIleIlePheProGln 420  
 Db 1201 ACTGTATAGGGGTCTGTGATCTCACTTGTGACCTTGTTCATCTGTATCTTCCCCCAG 1260  
 QY 421 LeuLeuCysValLeuPheIleTySgIyThrAsnThrTyGlyAlaValAlaIleGlyTyThr 440  
 Db 1261 CTGCTTGTGTACTCTTGTGTTAAGGAAACCAACACTATGGGGCCGTGGAGGTTATGTT 1320  
 QY 441 PheGlyLeuPheLeuArgIleThrGlyIyGlnProTyThrIleuGlnProLeuIle 460  
 Db 1321 TCTGGCTCTTCTGAGAAATACCTGAGAGGGAGCCATATCTGTATCTTACGCCCTTGATC 1380

QY 461 PheTyProGlyTyThrSerAspLysAsnGlyIleTyAsnGlnArgPheProPheIySgIn 480  
 Db 1381 TTCTATCCCTGGGCTATTAACCTGATGATATGATATATATATATATATATATATATAT 1440  
 QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValIleTyThrIleuAlaIySg 500  
 Db 1441 ACACTTGCATGATGATATCATATCTTCAACCAACATTTGATCTCTCATCTTACGCAAGTAT 1500  
 QY 501 LeuPheGlnSerGlyTyThrIleuProProLysLeuAspValPheAspAlaValAlaIySg 520  
 Db 1501 CTATTGAAAGTGAACCTTGGCCACTTAATATGATATATGATATGATCTGTGTGCAAG 1560  
 QY 521 HisSerGlyLysAsnMetAspLysThrIleLeuValArgAsnGlnLysAsnIleLysLeuAsn 540  
 Db 1561 CACAGTGAAGAAACATGATATAGACATTTCTTGTCAAAAATGAAATATTAATTAATAGAT 1620  
 QY 541 GlnLeuAlaProValIySgProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
 Db 1621 GAACCTTGACCTTGTGAAGCCAGACAGACATGACCCCTCAGCTCAACTTTCACCAATTA 1680  
 QY 561 GlnAlaLeuLeuAspValAspSerSerProGlyIySerGlyTyThrGlnAspAsnLeuGln 580  
 Db 1681 GAGGCTTCTCTGATGTATTCAGTCCAGAGAGGCTGTGGACTGAAGATATATTAACG 1740

RESULT 11  
 CQ722844 1743 bp DNA linear PAT 03-FEB-2004  
 LOCUS Sequence 8778 from Patent WO02068579.  
 DEFINITION CQ722844  
 ACCESSION CQ722844  
 VERSION CQ722844.1 GI:42283701  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 Kitz, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof  
 Patent: WO 02068579-A 8778 06-SEP-2002;  
 PE Corporation (NY) (US)  
 Location/Qualifiers  
 source 1..1743  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 1743  
 Score: 2795.00 Matches: 537  
 Percent Similarity: 96.6% Conservatve: 23  
 Best Local Similarity: 92.6% Mismatches: 20  
 Query Match: 93.4% Indels: 0  
 DB: 2 Gaps: 0

US-10-724-806-4 (1-580) x CQ722844 (1-1743)

QY 1 MetProPheHisValGlnIyLeuValAlaIleIleLeuPheTyThrLeuIlePheLeu 20  
 Db 1 ATGGCTTCCATGTGAAGAGCATGATAGCATGCTGTCTTACCTTCAATTTGCTG 60  
 QY 21 ValGlyIleThrAlaAlaTyThrLysAsnSerGlyAsnProGlnGlnArgSerGln 40  
 Db 61 GTTGAATATGGGCTGCTCGAGAAACCAAAACAGTGGAGCGCAGAAAGCGACGAA 120  
 QY 41 AlaIleIleValAlaGlyIyArgAspIleGlyLeuLeuValAlaGlyPheThrMetThrAla 60  
 Db 121 GCCATATAGTGGTGCGCGAGATATGTTGTTATGTTGTTGATTTACATGACAGCT 180  
 QY 61 ThrTyrValIyGlyIyTyThrIleAsnGlyIyThrAlaGlnAlaValTyThrIyProGlyCys 80







```

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrIleuIlePheLeu 20
DB 1 ATGGCTTTCCATGTGAAAGACTGATGCTATCATCTGTTCTTCACTTTTAATTTTGGCTG 60
QY 21 ValGlyIleTyrAlaIleTyrPheTyrIleuValAlaIleIleLeuPheTyrIleu 40
DB 61 GTTGGAAATATGGGCTGCTGAGAACCAAAACAGTGGCAGGACAGAGGCGACAGCAA 120
QY 41 AlaIleIleValGlyIleValArgAspIleGlyLeuLeuValGlyIlePheThrIle 60
DB 121 GCCATCATATGTTGGTGGCCAGATATTTGTTTATTTGGTGGATTTTCCATGACAGCT 180
QY 61 ThrTyrValGlyIleGlyTyrIleValIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB 181 ACCTGGGTGGAGAGGATATATCATATGACACAGCTGAAAGCAGTTATATGACAGGTTAT 240
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerIleuIleGlyIleu 100
DB 241 GGCTTAGCTTGGGCTCAGGACCAATTTGGATATTTCTTAGTCTGATTTTAAAGTGGCTG 300
QY 101 PhePheAlaIleValProMetIleArgSerIleGlyTyrValThrMetLeuAspPropheIleu 120
DB 301 TTCTTTGCAAAACCTATGCGTTCTMAAGGGGTATGTACCATGTTAGACCCGTTTCAGCAA 360
QY 121 IleTyrGlyIleValArgMetGlyIleLeuPheIleProAlaIleuMetGlyIleuMetPhe 140
DB 361 ATCTATGAAAGACATGGGCGGACTCTGTTTATTTCTGACCTGATGGAGAAATGTTTC 420
QY 141 TrpAlaAlaIleIlePheSerAlaIleuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGACGACATTTTCTGCTTTGGAGCCACATCATGATCATCATGATGAT 480
QY 161 ValAsnIleSerValIleValSerAlaIleuIleAlaIleuTyrThrIleuValGlyIle 180
DB 481 ATGCACATTTCTGTCATCTCTGACATCTGACATCTGACATCTGACATCTGACATCTG 540
QY 181 LeuTyrSerValAlaTyrThrAspValValGluIlePheCysIlePheIleGlyIleuTyr 200
DB 541 CTCTATCTGAGCTTACATGATGCTGATGCTGCTTTTGGATTTTGGAGGCTGATG 600
QY 201 IleSerValProPheAlaIleuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCAGGCTCCCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 221 HisAlaIleValTyrGlySerProTyrIleuGlyIleGluSerValGluValTyrThrTyr 240
DB 661 CATGCCAAATATACCAAAAGCCGTGGCGGAACTGTTGACTGATGAACTGATGACTT 720
QY 241 LeuAspAsnPheLeuIleuMetLeuGlyIleIleProTyrGluAlaTyrPheGluIleu 260
DB 721 CTGTATAGTTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGluValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTTCTCTCTCTCTCTCCACGACCACTATGCTCAAGTCTGCTTCTTGGCAGCTTTGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
DB 841 TCCCTGGTATGCGCCATCCAGCATCATATTTGGGCGCATTTGGACATCATCAGACTG 900
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrIleGluGluAlaAspMetIleu 320
DB 901 AACCAACTGCATATGAGGCTTCCAGATCCCAAGACTACAGAGAGGACAGATGATTTTA 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB 961 CCAATTTGTTCTGAGATCTCTGCTGCTGATATATTTCTTTGGTCTGGTGCAGTT 1020
QY 341 SerAlaIleValMetSerSerAlaAspSerSerIleuSerIleAspMetPheAla 360
DB 1021 TCTGCTGCTGTATATGATCATCAGCAATTTCTTCACTTGTGACAGAACTTCATGATTTGA 1080

```

```

QY 361 ArgAsnIleTyrGluIleuSerPheArgGluAsnAlaSerAspIleGluIleValTyrVal 380
DB 1081 CGGACATCTTACACACTTTCTTCCAGACAAATGCTTGGACAAAGAAATCGTTGGCTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
DB 1141 ATGCATATCAGATGTTTGTGTTTGGAGCATCTGCAACGCGCATGGCTTGTGTGACGAA 1200
QY 401 ThrValTyrGlyLeuTyrTyrIleuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1201 ACTGTATAGGCTCTGTTGATCTTCACTTGTGATGCTTGTATGATGCTTCTTCCCGC 1260
QY 421 LeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1261 CTGCTTTGTGATCTCTTTTAAAGGAACCAACACTTATGGGCGCTGACAGTATGTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyIleGluProTyrLeuTyrLeuGlnProLeuIle 460
DB 1321 TCTGGCTCTTCTCTGAGATTAAGTGGAGGAGCCATATCTGATCTTCAAGCCCTTGATC 1380
QY 461 PheTyrProGlyTyrTyrSerAspIleValIleTyrAsnGluIlePheProPheIle 480
DB 1381 TTCTACCTGCTATTTACCTGATGATATATGATATATATATATATATATATATATAT 1440
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
DB 1441 ACACCTTGCATGTTATCATCATTTCTTAAACCAACTTGGATCTTCCATCTAGCAAGTAT 1500
QY 501 LeuPheGluSerGlyThrLeuProProlysLeuAspValPheAspAlaValAlaArg 520
DB 1501 CTATTTGAAAGTGAACCTTGCACCTTAATATGATGATTTGATGCTGTTTGGCAAA 1560
QY 521 HisSerGluGluAsnMetAspIleThrIleLeuValArgAsnGluIleIleuAsn 540
DB 1561 CACAGTGAAGAAACATGATATGACATTTCTGCAAAATGAAATATATATATATATAT 1620
QY 541 GluLeuAlaProValIleProArgIleuSerLeuThrIleuSerSerThrPheThrAsn 560
DB 1621 GAACCTTGCACTGTGTAGACGACGACGACGACGACCTTACGCTCAACTTACCAATATA 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluIleGlySerGlyThrGluAspIleuGln 580
DB 1681 GAGGCTTCTTGTATGATTCAGTCCAGAAAGGCTGCGACATGATATATTTACAG 1740

```

RESULT 13  
 AR268949  
 LOCUS AR268949 1743 bp DNA linear PAT 10-APR-2003  
 DEFINITION Sequence 1 from patent US 6500643.  
 ACCESSION AR268949  
 VERSION AR268949.1 GI:23699686  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1743)  
 AUTHORS Wu D.-H., Gu Y., Millard W.J. and He Y.-J.  
 TITLE Human high affinity choline transporter  
 JOURNAL Patent: US 6500643-A 1 31-DEC-2002;  
 University of Florida; Gainesville, FL  
 FEATURES  
 source  
 1..1743  
 /organism="Unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 1743  
 Score: 2795.00 Matches: 537  
 Percent Similarity: 96.6% Conservative: 23  
 Best Local Similarity: 97.6% Mismatches: 20  
 Query Match: 93.4% Indels: 0  
 DB: 2 Gaps: 0

US-10-724-806-4 (1-580) x AR268949 (1-1743)

```

Oy 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTCCATGTCGAGGAGCATGATAGCTATCATCGTGTCTACTCTTATTTTGGCTG 60
Oy 21 ValGlyIleTPAlaAlaIleTPLeuThrLysAsnSerGlyAsnProGluLysSerGlu 40
Db 21 GTTGAAATATGGGCTGCTCGAGAACCAAAAACAGTGGGACGCGCAAGAGCGCGCA 120
Oy 41 AlaIleIleValGlyGlyYArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db 41 GCCATCATAGTTGGTGGCCGAGATATTGGTTTATGGTTGGTGAATTTACATGACAGCT 180
Oy 61 ThrTPValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 61 ACCTGGTCGAGGAGGATATCATGCAATGACAGCTGAAGCAAGTTTATGACAGTTAT 240
Oy 81 GlyLeuAlaIleTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleGluGlyLeu 100
Db 81 GGCTTAGCTTGGGCTCAGGACCAATGGATATCTCTTAGCTCATATTTAGTGGCCCTG 300
Oy 101 PhePheAlaLysProMetArgSerLysGlyTyrValIleThrMetLeuAspProPheLysGln 120
Db 301 TTCTTTGCAAAAACCTATGCTTCAAAAGGGGATATGACCATGTTAGACCCGTTTCA 360
Oy 121 IleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
Db 361 ATCTATGMAAAGCATGGGCGGACCTCGTTTATTCCTGACCTAGTGGAGAAATGTTTC 420
Oy 141 TPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGAGCAATTTCTCTGCTTTGGAGCCACATCAGCGTATCATCATGATGGAT 480
Oy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 481 ATGCACATTTCTGTATCATCTCTGCTCATCTGACATTCCTGTACACACTGGTGGAGGG 540
Oy 181 LeuTyrSerValAlaTyrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTyr 200
Db 541 CTCATCTCTGTCGCTACACTGATGCGTTCAGCTCTTTGCACTTTTGTAGGGCTGTGG 600
Oy 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
Db 601 ATCAGCGTCCCTTGGCATGTGCACATCTCGACGTGACAGACATGGGTTCACTCTGTG 660
Oy 221 HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGluValTyrThrTyr 240
Db 661 CATGCCAAATACCAAAAGCGGTGGCTGGGAACTGTGATCTCTGAAAGTCTACTCTTGG 720
Oy 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260
Db 721 CTGTATAGTTTCTGTGTGATGCTGGGTGAATCCCATGCAAGCATATTTTCAGAGG 780
Oy 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTCTCCAGCCACCATATGCTCAGAGCTGCTCTCTCTGCGACCTTTCGGG 840
Oy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
Db 841 TGCCGTGATGATGCCATCCCAACCATCTCATTTGGGGCATTTGGAGCCCTCCACAGACTGG 900
Oy 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
Db 901 AACCGACTGATATGGGCTTCCAGATCCCAAGACTACAGAGAGGACAGACATGATTTTA 960
Oy 321 ProIleValLeuGlnTyrLeuGlyProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATGTTCGACATATCTCTGCGCTGATATTTCTTTCTTTGGTCTGCTGACAGTT 1020
Oy 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGTGCTGTATATGTCATCAGACAGATTTCTTCATCTTGTACAGCAAGTTCCATTTTGA 1080

```

```

Oy 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValIleTPVal 380
Db 1081 CGGAACATCTACCAAGCTTCTCTTCAGACAAATGCTTCGCAAAAGAAATGTTGGGTT 1140
Oy 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGGCAATCACAGTGTGTGTGGAGCATCTCAACAGCCATGAGGCTTCTGACGAA 1200
Oy 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGATATGGGCTCTGCTGATCTCACTTGTATCATCTGATCTTCCCCAG 1260
Oy 421 LeuLeuCysValIlePheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTTGTATCTTCTTTAGGGAACCAACACTATGAGGCGCGTGGCAGGTTATGTT 1320
Oy 441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460
Db 1321 TCTGGCTCTTCTCGAGAAATACAGGAGGAGGACCATATCTGTATCTTACCCCTTGATC 1380
Oy 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db 1381 TTCTATCCCTGCTATTAACCCGATGATATGTTATATATATCAGAAATTTCCATTAAA 1440
Oy 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db 1441 ACACCTGCCATGATATCATCTTCTTAACCAACTTGGACATCTCTCATCTGACCAAGTAT 1500
Oy 501 LeuPheGlySerGlyTyrLeuProProLysLeuAspValPheAspAlaValValAlaArg 520
Db 1501 CTATTTGAAATGGAACCTTCCACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1560
Oy 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATATGACATCTTGTGCAAAATGAAAAATTAATTAAT 1620
Oy 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGACATTTGGAACCCAGACAGACATGACCCCTCACTTCAACCAATTA 1680
Oy 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyTyrGluAspAsnLeuGln 580
Db 1681 GAGGCTTCTTCTGATGTATTCAGTCCAGAGAGGCTTGGAGACTGAAGATTAATTA 1740

```

RESULT 14  
AF276871 1743 bp mRNA linear PRI 27-NOV-2000  
LOCUS AF276871  
DEFINITION Homo sapiens high affinity choline transporter (SLC5A7) mRNA.  
ACCESSION AF276871  
VERSION AF276871.1 GI:10998441  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homosapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 1743)  
Apparundaram,S., Ferguson,S.M., George,A.L., Jr. and Blakely,R.D.  
Molecular cloning of a human, hemicholinium-3-sensitive choline  
transporter  
Biochem. Biophys. Res. Commun. 276 (3), 862-867 (2000)  
JOURNAL PUBMED 11027560  
REFERENCE 2 (bases 1 to 1743)  
Apparundaram,S., Ferguson,S.M. and Blakely,R.D.  
AUTHORS Direct Submission  
TITLE Submitted (03-JUN-2000) Department of Pharmacology and Center for  
Molecular Neuroscience, Vanderbilt University, 23rd Avenue South at  
Pierce, Nashville, TN 37232-6420, USA  
JOURNAL  
FEATURES  
source  
1. 1743  
/organism="Homo sapiens"

gene  
CDS  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2q12"  
1..1743  
/gene="SLC5A7"  
1..1743  
/note="hCMT: solute carrier family 5 member 7"  
/product="high affinity choline transporter"  
/protein\_id="AAG25940.1"  
/db\_xref="GI:10998442"  
/translation="MAFVEGLIAIVFLLILVGMAMRTKSGAERSEAVIV  
GGRDGLVGFMTATWVGGINGTAEVAVVPGVGLMAQAQPIGSLILGLIEF  
AKPMRSKGVYTMIDPQQIYKGRMGLRPLMGMFMAAIFSLKGTISVITVD  
MHSVILSLALITLVGLISVATDVQLFCIFVGLISVPLSHPAVADIGFT  
AVHAKYQKPMIGTVDSEVSWLDSFLMLDGIQWAFQFRLSSSSATVADISFL  
AAGGLVMAIPAILIGALIGASTDMNQAYGLPDKTBEDMILPLVLCYLPVYSF  
FELGASAVMSADSSILSASSMEFANITYLSFROMADKEIVWMRTVFEFGASA  
TAMALLTKTVGLAMYSPLVIVIPROLICVLPKGTNTVAGVAGVGLRITGG  
EYVLIQPLIFRYGYPDNGIYNQKPKITLAMYISPLFNICISLALYRESGLP  
PLDVPDAVAVARSHENMDKTLIVKNENIKLDELAVKPRQSWTSTITNKAFLVD  
DSSPDSGTEHDNIQ"

ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 1743  
Score: 2795.00 Matches: 537  
Percent Similarity: 96.6% Conservative: 23  
Best Local Similarity: 92.6% Mismatches: 20  
Query Match: 93.4% Indels: 0  
DB: Gaps: 0

US-10-724-806-4 (1-580) x AF276871 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaAlaIleLeuPheTyreuleuIlePheLeu 20  
DB 1 ATGGCTTTCATGTGAGAGACTGATGATCATGTGTTCACTTCAATTTGGCTG 60  
QY 21 ValGlyIleThrAlaAlaThrIlySerIlyAsnSerGlyAsnProGluGluArgSerGlu 40  
DB 61 GTTGGAAATATGGGCTCTCGAGAACCAAAACAGTGCAGCCGACAGAGCCGACGAA 120  
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyIlyPheThrMetThrAla 60  
DB 121 GCCATCATATGTTGGTGGCCGAGATATTTGTTATTTGGTGGATTTACCATGACAGCT 180  
QY 61 ThrTrpValGlyGlyGlyIlyLeaenGlyThrAlaGluAlaValTyrglyProGlyCys 80  
DB 181 ACCTGGGTGGAGAGAGGTATATCATGACACAGCTGAAAGCACTTATATGACAGCTTAT 240  
QY 81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyreuleuSerIleuIleuGlyIleu 100  
DB 241 GGCTTAGCTTGGGCTCAGGACCAACATTTGATATTTCTTAGTCTGATTTAGGTCGCTG 300  
QY 101 PhePheAlaIlyProPheArgSerIlySerIlyValThrMetLeuAspProPheIlyGln 120  
DB 301 TTCTTTGCAAAACCTATGCGTTCAAGGGGTATGTACACATGTTAGACCCGTTTCAGCA 360  
QY 121 IleTyrglyIlyAsnArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyIleuMetPhe 140  
DB 361 ACTTATGAGAAACGATGGGCGGACTCTGTTTATTTCTGCACTGATGGAGAAATGTTTC 420  
QY 141 ThrPheAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
DB 421 TGGGCTGACAGCAATTTCTCTGCTTGGAGCCACCATCAGCGGTGATCGATGAGAT 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleuTyrrhrLeuValGlyIly 180  
DB 481 ATGCACATTTCTCTCATCTCTGCACTCATTTGCCACTCTGACACACTGGTGGAGGG 540  
QY 181 LeuTyreSerValAlaTyrrhrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200

DB 541 CTCTATTTCTGGGCTTACACTGATGCTTACGCTTTTGGATTTTGTAGGGCTGTGG 600  
QY 201 IleserValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
DB 601 ATCAGGCTCCCTTGGCATATGTCATCTCGAGTCGACAGATCGGTTCACTGCTGTG 660  
QY 221 HisAlaIlyTyrglnSerProThrIleuGlyIlyThrIleGlySerValGluValTyrrhrTrp 240  
DB 661 CATGCCAATTACCAAAAGCGGTGGCTGGGACCTGTTGACTCATCTCACTTCTTGG 720  
QY 241 LeuAspAsnPheLeuLeuLeuMetIleuGlyIlyIleProTrpGlnAlaTyrrhrPheGlnArg 260  
DB 721 CTGTATAGTTTCTGTTGTTGATGCTGGGTGGAATCCATGCAACCATATCTTTCAGAG 780  
QY 261 ValLeuSerSerSerSerAlaThrTyrrhArgIleValLeuSerPheLeuAlaIlePheGly 280  
DB 781 GTTCTCTCTTCTTCTCCAGCCACTATGCTCAAGTGTCTCTTCTGCGAGCTTTCGG 840  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaIleSerThrAspTrp 300  
DB 841 TGCCTGTGATAGGCAATCCAGCCATATCTATTTGGGCCATTTGAGCATCAACAGCTGG 900  
QY 301 AsnGlnThrAlaTyrglyTyrrhrProAspProIlySerIlyGluGluAlaAspMetIleu 320  
DB 901 AACCAAGATGCATATGGGCTTCCAGATCCCAAGACTACAGAGAGGACACATGATTTTA 960  
QY 321 ProIleValLeuGlnTyrrhrLeuCysProValTyrrhrIleSerPhePheGlyIleuGlyAlaVal 340  
DB 961 CCATATGTTCTCAGATATCTGCTGCTGATATTTCTTGGTCTTGGTGGAGTT 1020  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
DB 1021 TCTGCTGCTGTTATGTCATCAGCAGATTTCTTCATCTTTGTCAGCAAGTTCCAGTTTGA 1080  
QY 361 ArgAsnIleTyrglnLeuSerPheArgGlnAsnAlaSerAspIlyGluIleValIleThrVal 380  
DB 1081 CGGAACATCTACACAGCTTCTTCAGACAAATATGCTTCGACAAAGAAATCGTTGGGTT 1140  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIly 400  
DB 1141 ATGCCAATCAGATGTTTGTGTTGGAGCATTCGCAACGCCATGCGCTTGTGCAGAA 1200  
QY 401 ThrValTyrglyIlyLeuTrpTyrrhrLeuSerSerAspLeuValTyrrhrIleIlePheProGln 420  
DB 1201 ACTGTATAGGCTCTGAGTACTGACTGCTGACTTGTATACATCGTATCTTCCCCAG 1260  
QY 421 LeuLeuCysValLeuPheIleIlyGlyIlyThrAsnThrTyrglyAlaValAlaGlyTyrrhr 440  
DB 1261 CTGCTTGTGTACTTGTGTTAAGGGAACAACACTATGGGCGCGTGGCAGGTTATGTT 1320  
QY 441 PheGlyLeuPheLeuArgIleThrGlyIlyGluProTyrrhrLeuTyrrhrLeuGlnProLeuIle 460  
DB 1321 TCTGGCTCTCTCTGGAATTAAGTGAAGGAGGCCATATCTGTATCTTACAGCCCTTATC 1380  
QY 461 PheTyrrhrProGlyTyrrhrSerAspIlyAsnGlyIleTyrrhrAsnGlnArgPheProPheIly 480  
DB 1381 TTCTACCTGCTATTAACCTGATGATATGATATATATATATATATATATATATATAT 1440  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrrhrLeuAlaIlyTyrrhr 500  
DB 1441 ACACCTTGCATGTTATCATCTTAAACAACATTTGATCTCCATATCAGCAAGTAT 1500  
QY 501 LeuPheGlnSerGlyTyrrhrLeuProPheIlyAspValPheAspAlaValAlaIlyAsp 520  
DB 1501 CTATTTGAAAGTGAACCTTGCACCTTAATTAAGTATTTGATCTGTTGTTGCAAG 1560  
QY 521 HisSerGlnGluAsnMetAspIlyThrIleLeuValArgAsnGluAsnIleIlyLeuAsn 540  
DB 1561 CACAGTGAAGAAACATGATTAAGCAATTTCTGTCAAAATATGAATTAATTAATTAAGAT 1620  
QY 541 GluLeuAlaProValIlyProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIly 560

Db 1621 GAAGTTCAGTTCGACGACGACAGACATGACCTCAGCTCAACTTACCAATATA 1680  
Qy 561 GluAlaLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
Db 1661 GAGGCTTCCTTGATGATTCAGTCCAGTCCAGAAAGGCTCGGACATGAGATTAATTACAG 1740  
RESULT 15  
LOCUS HSA401466 1813 bp mRNA linear PRI 16-AUG-2000  
DEFINITION Homo sapiens mRNA for high affinity choline transporter (CHT1 gene).  
ACCESSION AJ401466  
VERSION AJ401466.1 GI:9843753  
KEYWORDS CHT1 gene; high affinity choline transporter.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Wieland,A., Bonisch,H. and Bruns,M.  
TITLE Molecular cloning of the human and murine high affinity choline transporters and characterization of the human gene-structure  
REFERENCE 2 (bases 1 to 1813)  
AUTHORS Bruns,M.  
TITLE Direct Submission  
JOURNAL Submitted (14-AUG-2000) Bruns M., University of Bonn, Pharmacology and Toxicology, Reuter str. 2b, D-53113 Bonn, GERMANY  
FEATURES  
source  
1.1813  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2q11-13"  
/tissue\_type="hypothalamus"  
1.1813  
/gene="CHT1"  
19.11761  
/function="sodium- and chloride-dependent reuptake of choline"  
/experiment="experimental evidence, no additional details recorded"  
/codon\_start=1  
/product="high affinity choline transporter"  
/protein\_id="CAC03717.1"  
/db\_xref="GI:9843754"  
/db\_xref="GDB:984375"  
/translation="MAFHEGLIAIVLYLLILVLGVIAAMETKNSGSAERSEAIIV  
GARDIGLVGGFTMTATVGGVINGTAEAAYVPGYLAQAQAGYSLILGLGF  
AKPMRSKGVMTLDPFOIYGRMGLEFIPALMGEMWAAIFPALGATTSVILDVD  
MHISYISALIAATLYLGLSVATVVOVFCIFGLMISVFPALSHPAVADIGFT  
AVHAKYOKPMICTUDNSSEVSWLDSFILMIGITPMQVAFQVIVSSSRTAQTUSPL  
AARGLVNAIPAILIGATGASTDMQTAIGLPDKTEBADLITVLQYLDQVYTSF  
FGLGAVSAVMSSADSSILSASMPARNIYQSFQNSDKEIVWMRIYVFFGASA  
TAMALLFTKVGLSSDLVYVIFPOLCVLPKGTNYGAVAGVSGFLRTGG  
EPYVLQPLIFPGYYPDDNGIYNOKFPFKTLAMYSFLTNICSYLAKYLESGTLP  
PKLDVFDLVAVARHSEEMDKTILVNENIKDELALVVRQSMTLISPTFNKEAVLDV  
DSSPEGSSTEDNLQ"

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 1813  
Score: 2795.00 Matches: 537  
Percent Similarity: 96.6% Conservative: 23  
Best Local Similarity: 92.6% Mismatches: 20  
Query Match: 93.4% Indels: 0  
Gaps: 0  
US-10-724-806-4 (1-580) x HSA401466 (1-1813)

Qy 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
Db 19 ATGGCTTCCATGTGGAAGACTAGTACATCGCTTCACTCAATTTTGGCTG 78  
Qy 21 ValGlyIleTPrAlaIleTPrIleThrIleAsnSerGlyAsnProGluGluAspSerGlu 40  
Db 79 GTTGGAAATATGGCTGCTCGGAGAACCAAAACAGTGGACGACGAGAGACGACGAA 138  
Qy 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
Db 139 GCCATCATAGTTGGTGGCCGAGATATTCGTTATTTGGTGGATTAACATGACACT 198  
Qy 61 ThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
Db 199 ACCGTGGTCGAGGAGGGGTATATCATGTCAGCTGAAAGCATTTATGTCACAGTTAT 258  
Qy 81 GlyLeuAlaTPrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
Db 259 GGCTTAGCTGGGCTCAGGACCAATGGATATTCCTTAGCTGATTTAGTGCGCTG 318  
Qy 101 PhePheAlaIleProMetArgSerIleGlyTyrValThrMetLeuAspProPheIleGln 120  
Db 319 TTCTTTGCAAAACCTATCGCTTCAAGGGGTATGACCATGTTAGACCCGTTTCAGCA 378  
Qy 121 IleTyrGlyIleArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
Db 379 ATCTATGGAAGAAACGATGGGCGGACTCGCTTTATTCCTGACATGATGGAGAAATGTTTC 438  
Qy 141 ThrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
Db 439 TGGCTGACGAAATTTTCTGCTTTGGAGCCACCATCAACGGATATGATGATGAT 498  
Qy 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
Db 499 ATGCACATTTCTGTCATCATCTGTCACATTCGACATTCGACACTGTCACACTGGTGGAG 558  
Qy 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200  
Db 559 CTCTATTCCTGTCGCTTACATGATGTCCTGAGCTCTTTTTCATTTTGGAGGCTGAG 618  
Qy 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
Db 619 ATCAGCTGCCCTTTGATGTCATCTGTCACATCTGTCACATGTCAGATGCTGCTGTG 678  
Qy 221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrThrTrp 240  
Db 679 CATGCCAATATCAAAACCGGCTGGGAGACTGTGACTCATCTGAGAGTCACTCTTGG 738  
Qy 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArg 260  
Db 739 CTGTATATCTTCTGTTTGTATGCTGGGTGGAAACCCATGCGCAAGCACTTCTGAGGG 798  
Qy 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280  
Db 799 GTTCTCTCTTCTTCTCTGACCACTATGCTCAAGTGCCTGCTTCTGACGCTTTCGGG 858  
Qy 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300  
Db 859 TGCGTGGATGAGCCATCCAGCATACTCATTTGGGGCATTGGAGCATCAACAGACTGG 918  
Qy 301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrIleGlyGluAlaAspMetIleLeu 320  
Db 919 AACCAAGCTGATATAGGGCTTCCAGATCCCAAGACTACAGAAAGGCGACATGATTTTA 978  
Qy 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
Db 979 CCAATGTCTTCCAGATATCTGCTGCTGTATATTTTCTTGTGCTGCTGCTGCTGCT 1038  
Qy 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
Db 1039 TCTGCTGCTGTATATGATCATGACAGATTTCTTCATCTTGTGACGACAGATGCTTGTGCA 1098

QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpVal 380  
|||  
DB 1099 CGGAACATCTACACAGCTTCTTCAGACAAATCGTTGCGACAAAGAAATCGTTGGGTT 1158  
|||  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLys 400  
|||  
DB 1159 ATCCGAATCACAGTGTGGTTGGAGCATCTGCAACAGCCATGGCGCTTGCTGACGAAA 1218  
|||  
QY 401 ThrValTyrGlyLeuTrpPheLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
|||  
DB 1219 ACTGTATAGGCTCTGGTACCTCAGTTCGACCTGTGTACATCGTTATCTTCCCCAG 1278  
|||  
QY 421 LeuLeuCyValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
|||  
DB 1279 CTGCTTGTGTACTCTTTGTTAAGGACCAACACCATATGGGCGCGGCGAGGTATGTT 1338  
|||  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460  
|||  
DB 1339 TCTGGCCTCTCTCGAATAAATCTGAGGGAGCCATATCTGTATCTTCAGCCCTTGATC 1398  
|||  
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
|||  
DB 1399 TTCTACCTGGCTATTACCTGATGATATATGATATATATATATATATATATATATAT 1458  
|||  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCyValSerTyrLeuAlaLysTyr 500  
|||  
DB 1459 ACACTTGCCTAGTGTATCATATTCTTAACCAACATTTGCATCTCTATCTAGCCAAATAT 1518  
|||  
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520  
|||  
DB 1519 CTATTGAAAGTGAACCTTGCCACCTAAATTAATGATGATGATGATGATGATGATGATG 1578  
|||  
QY 521 HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn 540  
|||  
DB 1579 CACAGTGAAGAAAAACATGATTAAGACATCTTGTCAAAAATGAAATATTAATTAATGAT 1638  
|||  
QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
|||  
DB 1639 GAACTTGCACTTGTGAAGCCACGACGACGATGACCTCAGCTCACTTCAACCAATAA 1698  
|||  
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
|||  
DB 1699 GAGGCTTCTCTTGATGTGATTCAGTCCAGAAAGGCTGTGGACTGAAAGATTAATTAACAG 1758  
|||

Search completed: July 11, 2006, 02:03:20  
Job time : 9566 secs

**This Page Blank (uspio)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:25 ; Search time 1047 Seconds  
(without alignments)  
11607.096 Million cell updates/sec

Title: US-10-724-806-3

Sequence: 1 atgccttcacatgctggaag.....ctgaagaataatcacaatga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_8:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*  
15: geneseqn2006s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1743	100.0	1743	10	ADD50640	ADD50640 cDNA enco
2	1743	100.0	1743	10	ADD50660	ADD50660 cDNA enco
3	1743	100.0	1743	14	ADV77898	ADV77898 Mouse hig
4	1743	100.0	1743	14	ADV77918	ADV77918 Mouse hig
5	1730.2	99.3	1743	4	AAFB1713	AAFB1713 Mouse hig
6	1730.2	99.3	1743	4	AAAD02457	AAAD02457 Mouse P4P
7	1616.6	92.7	1743	4	AAFB1711	AAFB1711 Rat high
8	1616.6	92.7	4904	10	ADD50642	ADD50642 cDNA enco
9	1616.6	92.7	4904	14	ADV77900	ADV77900 Rat high
10	1378.2	79.1	1743	9	ABX94338	ABX94338 Human CDN
11	1375	78.9	1743	4	AAFB1712	AAFB1712 Human hig
12	1375	78.9	1743	5	AAHA9207	AAHA9207 Human CHO
13	1375	78.9	1743	10	ADD50638	ADD50638 cDNA enco
14	1375	78.9	1743	14	ADV77896	ADV77896 Human hig
15	1375	78.9	1813	14	ADD50646	ADD50646 High-affi
16	1375	78.9	1813	14	ADV77904	ADV77904 Human hig
17	502.8	28.8	8760	5	AAHA9206	AAHA9206 Human CHO
18	502.8	28.8	26323	14	ABEB1646	ABEB1646 Human sci

19	502.8	28.8	119040	10	ADD50656	ADD50656 BAC seque
20	502.8	28.8	119040	14	ADV77914	ADV77914 Human BAC
21	502.8	28.8	142289	10	ADD50651	ADD50651 BAC seque
22	502.8	28.8	142289	14	ADV77909	ADV77909 Human BAC
23	388.8	22.3	1833	12	ADJ93831	ADJ93831 cDNA enco
24	356.4	20.4	1731	386	AAAD02461	AAAD02461 Mouse P4P
25	355.8	20.4	1731	4	AAFB1710	AAFB1710 C. elegans
26	355.8	20.4	1985	10	ADD50644	ADD50644 High-affi
27	355.8	20.4	1985	14	ADV77902	ADV77902 Nematode
28	321	18.4	329	12	ADP08298	ADP08298 Subacute
29	278.8	16.0	1729	4	ABL29569	ABL29569 Drosophila
30	237.6	13.6	1461	6	ABK73210	ABK73210 Bacillus
31	178	10.2	4223	4	ABL29568	ABL29568 Drosophila
32	145.6	8.4	10140	5	AAHA9201	AAHA9201 Human CHO
33	145.6	8.4	119040	10	ADD50656	ADD50656 BAC seque
34	145.6	8.4	119040	14	ADV77914	ADV77914 Human BAC
35	145.6	8.4	142289	10	ADD50651	ADD50651 BAC seque
36	145.6	8.4	142289	14	ADV77909	ADV77909 Human BAC
37	132.4	7.6	455	4	AAI11913	AAI11913 Probe #18
38	132.4	7.6	455	4	ABA53620	ABA53620 Human fce
39	132.4	7.6	455	4	AAI33245	AAI33245 Probe #18
40	132.4	7.6	455	4	ABA23372	ABA23372 Probe #18
41	132.4	7.6	455	4	AAK27341	AAK27341 Human bon
42	132.4	7.6	455	4	AAK01883	AAK01883 Human bira
43	132.4	7.6	455	4	ABS26913	ABS26913 Human liv
44	132.4	7.6	455	5	AAI01851	AAI01851 Probe #18
45	132.4	7.6	1094	5	AA882193	AA882193 DNA encod

## ALIGNMENTS

RESULT 1	
ADD50640	
ID	ADD50640 standard; cDNA; 1743 BP.
XX	
XX	ADD50640;
AC	
DT	15-JUN-2004 (first entry)
XX	
DE	cDNA encoding mouse high-affinity choline transporter (mCHT) #1.
XX	
KM	Mouse; high-affinity choline transporter; mCHT; cholinergic function;
KM	Parkinson's disease; Huntington's disease; Alzheimer's disease;
KM	schizophrenia; dysautonomia; myasthenia gravis; brain;
KM	cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;
XX	neuroprotective; neuroleptic; gene; ss.
OS	Mus sp.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..1743
FT	/*tag= a
FT	/product= "mCHT #1"
XX	
PN	US2003114399-A1.
XX	
PD	19-JUN-2003.
XX	
PF	23-JUL-2001; 2001US-00911077.
XX	
PR	23-JUL-2001; 2001US-00911077.
XX	
PA	(BLAK/) BLAKELY R D.
PA	(APPA/) APPARUNDARAM S.
PA	(FERG/) FERGUSON S.
XX	
PI	Blakely RD, Apparsundaram S, Ferguson S;
XX	
DR	WPI; 2003-810914/76.
XX	P-Psdb; ADD50641.
PT	Novel isolated polynucleotide encoding human or mouse high affinity

PT choline transporter polypeptide, useful in gene therapy to increase  
PT cholinergic function in a cell of a patient suffering from Alzheimer's  
PT disease.

XX Claim 30; SEQ ID NO 3; 74pp; English.

CC The present invention relates to the isolation of polynucleotide  
CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic  
CC function in the cell that is in a patient having Parkinson's disease,  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signaling. The present sequence encodes mCHT. Note: The sequence data  
CC for this patent was obtained in electronic format directly from the USPTO  
CC web site at seqdata.uspto.gov.

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

Query Match 100.0%; Score 1743; DB 10; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTTCCATGTTGGAAGAGCTGTGATGCTATATCTTTTACCTCTTATTTTCTG 60  
DB 1 ATGCTTTCCATGTTGGAAGAGCTGTGATGCTATATCTTTTACCTCTTATTTTCTG 60  
QY 61 GTTGAATATGAGGCTGCATGAGAAAACAAAACAGCGGACCCAGAGAGCGCAGTGA 120  
DB 61 GTTGAATATGAGGCTGCATGAGAAAACAAAACAGCGGACCCAGAGAGCGCAGTGA 120  
QY 121 GCCATCATAGTCGGGGGCGGTGACATGTTGTTGTTGTTGTTTACCATGACAGCC 180  
DB 121 GCCATCATAGTCGGGGGCGGTGACATGTTGTTGTTGTTGTTTACCATGACAGCC 180  
QY 181 ACCTGGGTTGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 181 ACCTGGGTTGAGAGAGGCTACATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 241 GGTCTAGCTTGGGCTCATGACACCATTTGATTTCTCTGAGTCTAATTTTAACTG 300  
DB 241 GGTCTAGCTTGGGCTCATGACACCATTTGATTTCTCTGAGTCTAATTTTAACTG 300  
QY 301 TTTTTCGGAACCTATGCGTTCCAAAGGATATGATGATGTTGAACCATTTCAAACG 360  
DB 301 TTTTTCGGAACCTATGCGTTCCAAAGGATATGATGATGTTGAACCATTTCAAACG 360  
QY 361 ATCTATGGAAGAGCGATGGGTGGGCTGCTTTTCATCCCTGACATGATGGAGAGAT 420  
DB 361 ATCTATGGAAGAGCGATGGGTGGGCTGCTTTTCATCCCTGACATGATGGAGAGAT 420  
QY 421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGGGCCACATAGCGTATCTTATATGAT 480  
DB 421 TGGGCTGAGCAATTTTCTCTGCAATTAAGGGGCCACATAGCGTATCTTATATGAT 480  
QY 481 GTGAACATATGGGCTATGTTCTCTGACATCATTTGGCATTCTTTAATCCCTAATG 540  
DB 481 GTGAACATATGGGCTATGTTCTCTGACATCATTTGGCATTCTTTAATCCCTAATG 540  
QY 541 CTCTACTCTGTCATATATGATGATGTTGTCAGCATTTCTGATTTTATAGAGCTG 600  
DB 541 CTCTACTCTGTCATATATGATGATGTTGTCAGCATTTCTGATTTTATAGAGCTG 600  
QY 601 ATCACTGTCCTTTTGGCTGTCATCTCTGAGATCAGGACATGAGATTCACAGCTG 660  
DB 601 ATCACTGTCCTTTTGGCTGTCATCTCTGAGATCAGGACATGAGATTCACAGCTG 660  
QY 661 CATGTAATATCCAGAGTCCCTGGCTGGGACCATTTGAATCAGTTGAAGTTCACCT 720

DB 661 CATGTAATATCCAGAGTCCCTGGCTGGGACCATTTGAATCAGTTGAAGTTCACCT 720  
QY 721 CTGTAATATTTTCTGTTATATGATGCTGGGTGAATCCCATGGCAAGCTTCCAGAG 780  
DB 721 CTGTAATATTTTCTGTTATATGATGCTGGGTGAATCCCATGGCAAGCTTCCAGAG 780  
QY 781 GTCTCTTATATCTCAGGACCATGATGCTAGATCTGCTTCTTCTGAGCTTTTGG 840  
DB 781 GTCTCTTATATCTCAGGACCATGATGCTAGATCTGCTTCTTCTGAGCTTTTGG 840  
QY 841 TGGCTGGATGAGGCTCTACCCGCCATATGATGATGAGAGCTATTTGAGCTTCCAGAG 900  
DB 841 TGGCTGGATGAGGCTCTACCCGCCATATGATGATGAGAGCTATTTGAGCTTCCAGAG 900  
QY 901 AACGAGATCTGCTACGGGTATCCAGATCCAGACTAAGAGAGAGAGAGAGATGCTC 960  
DB 901 AACGAGATCTGCTACGGGTATCCAGATCCAGACTAAGAGAGAGAGAGAGATGCTC 960  
QY 961 CCGATCTTTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 CCGATCTTTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 TCAGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 1021 TCAGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1081 CGGAATATCTACAGCTTCTCTTCCAGCAAAATGATGATGATGATGATGATGATGAT 1140  
DB 1081 CGGAATATCTACAGCTTCTCTTCCAGCAAAATGATGATGATGATGATGATGATGAT 1140  
QY 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 ACTGCTATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 1201 ACTGCTATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
QY 1261 CTGCTCTGCTGATCTTCTTCAATGAAGAACCAACTTATGAGGAGCTGCTGCTTAT 1320  
DB 1261 CTGCTCTGCTGATCTTCTTCAATGAAGAACCAACTTATGAGGAGCTGCTGCTTAT 1320  
QY 1321 TTTGAGATCTTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
DB 1321 TTTGAGATCTTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
QY 1381 TTTTACCTGCTTATTTACTCTGACAAAGATGATATCAATCAAGAGCTTCCATTTAA 1440  
DB 1381 TTTTACCTGCTTATTTACTCTGACAAAGATGATATCAATCAAGAGCTTCCATTTAA 1440  
QY 1441 ACTCTCTCATGCTTACCTCATTTCTTCAACCAATTTGCTTCTTATCTAGCCAGAT 1500  
DB 1441 ACTCTCTCATGCTTACCTCATTTCTTCAACCAATTTGCTTCTTATCTAGCCAGAT 1500  
QY 1501 CTATTTGAAGGAACTTCTGCTCAAAATTAAGTATGATTTGATGCTGCTGCTGCAAG 1560  
DB 1501 CTATTTGAAGGAACTTCTGCTCAAAATTAAGTATGATTTGATGCTGCTGCTGCAAG 1560  
QY 1561 CACAGTGAAGAGACATGAGACAGACATTTAGTCAAGAAATGAAATATCAATTAAT 1620  
DB 1561 CACAGTGAAGAGACATGAGACAGACATTTAGTCAAGAAATGAAATATCAATTAAT 1620  
QY 1621 GAATCTGACCTGTCGAAACCTGCGAGAGAGCTTAACCTTCACTTCAACCAATTAAG 1680  
DB 1621 GAATCTGACCTGTCGAAACCTGCGAGAGAGCTTAACCTTCACTTCAACCAATTAAG 1680  
QY 1681 GAGGCTCTCTGATGATTTTCAAGTCCGAGGAGGCTGAGGAGTCAAGATTAATTA 1740  
DB 1681 GAGGCTCTCTGATGATTTTCAAGTCCGAGGAGGCTGAGGAGTCAAGATTAATTA 1740  
QY 1741 TGA 1743





Db		CGGAATATCTACAGACTTTCCTCCAGACAAATGATCAGACAAGAAATTGTGTGG99TC	1140
Oy	1141	ATGAGGATCATCTGTGCTTGTTGTGGAGACATCTGCACAGCCATGCGTTTGCTGACGAG	1200
Db	1141	ATGAGGATCATCTGTGCTTGTTGTGGAGACATCTGCACAGCCATGCGTTTGCTGACGAG	1200
Oy	1201	ACTGTGATGAGGCTCTGGTACCTTGAGCTCAACCCTGTGCTAATCATCATCTTCCCACAG	1260
Db	1201	ACTGTGATGAGGCTCTGGTACCTTGAGCTCAACCCTGTGCTAATCATCATCTTCCCACAG	1260
Oy	1261	CTGCTCTGTGTACTCTTTCATCAAAGAAACCAACTTAATGGGGCAGTTGCTGTTTAAAT	1320
Db	1261	CTGCTCTGTGTACTCTTTCATCAAAGAAACCAACTTAATGGGGCAGTTGCTGTTTAAAT	1320
Oy	1321	TTTGGACTATTCCTGAGAAATTAACGGAGAGAGCCATATCTATACTTGACAGCCCTTAATC	1380
Db	1321	TTTGGACTATTCCTGAGAAATTAACGGAGAGAGCCATATCTATACTTGACAGCCCTTAATC	1380
Oy	1381	TTCTAACCTGTGTTATTTACTCTGACAGAATGGTATATACAATCAAGAGTTGCCATTAAA	1440
Db	1381	TTCTAACCTGTGTTATTTACTCTGACAGAATGGTATATACAATCAAGAGTTGCCATTAAA	1440
Oy	1441	ACTCTCTGCATGCTTACCTCAATCTTTTCCAAACATTTGTGTTTTCTTATCTTAGCCAAGAT	1500
Db	1441	ACTCTCTGCATGCTTACCTCAATCTTTTCCAAACATTTGTGTTTTCTTATCTTAGCCAAGAT	1500
Oy	1501	CTATTTGAAAGTGGAAACCTTGCTCCAAATAATAGATGATTTGATNGCTGTGTGCGAAG	1560
Db	1501	CTATTTGAAAGTGGAAACCTTGCTCCAAATAATAGATGATTTGATNGCTGTGTGCGAAG	1560
Oy	1561	CACAGTGAAAGAGACATGAGCAAGACCTTTCTAGTCAGAAAAATGAANAATTAATTAAT	1620
Db	1561	CACAGTGAAAGAGACATGAGCAAGACCTTTCTAGTCAGAAAAATGAANAATTAATTAAT	1620
Oy	1621	GAACTGTGACCTGTGTAACCTCGGAGAGAGCCTTAACCTCAGTTCACAATCACCAGTAAG	1680
Db	1621	GAACTGTGACCTGTGTAACCTCGGAGAGAGCCTTAACCTCAGTTCACAATCACCAGTAAG	1680
Oy	1681	GAGGCCCTCCTTGATGTTGATTCAGTCCGAGAGGGCTGCGAGACTGAAGATTAATTCAA	1740
Db	1681	GAGGCCCTCCTTGATGTTGATTCAGTCCGAGAGGGCTGCGAGACTGAAGATTAATTCAA	1740
Oy	1741	TGA 1743	
Db	1741	TGA 1743	
<hr/>			
RESULT 3			
ID	ADV77898	standard; cDNA, 1743 BP.	
XX	ADV77898;		
AC			
XX	24-FEB-2005 (first entry)		
DT			
XX	Mouse high affinity choline transporter (mCHT) cDNA.		
DE			
XX			
XX	Choline transporter; neuromuscular disorder;		
KM	autonomic nervous system disorder; central nervous system disorder;		
KW	Parkinsons disease; Huntingtons chorea; genetic disorder;		
KV	Alzheimers disease; degeneration; neurological disease; schizophrenia;		
KW	psychiatric disease; myasthenia gravis; immune disorder; gene therapy;		
KM	cns-gen.; noctropic; neuroprotective; antiparkinsonian; anticonvulsant;		
KX	mucular-gen.; neuroleptic; DNA purification; gene; ss.		
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1743	
FT		/+tag= a	
XX		/product= "Mouse CHT protein"	
PN	US2004248838-A1.		

XX	09-DEC-2004.
PD	
XX	
XX	01-DEC-2003; 2003US-00724806.
PF	
XX	
PR	23-JUL-2001; 2001US-00911077.
XX	
PA	(UYVA-) UNIV VANDERBILT.
XX	
Pt	Blakely RD, Apparsundaram S, Ferguson S;
XX	
DR	WPI; 2005-020586/02.
XX	
DR	P-P5DB; ADV77899.
XX	
DR	GENDANK; AF276872.
XX	
PT	New human high affinity choline transporter (CHT) cDNA, useful for
PT	treating a neuromuscular, autonomic or central nervous system disorder,
PT	including Parkinson's disease, Huntington's disease, Alzheimer's,
PT	schizophrenia.
XX	
PS	Claim 30; SEQ ID NO 3; 73bp; English.
XX	
CC	The present invention provides polynucleotides encoding novel high
CC	affinity choline transporters (CHTs), methods for their use in screening
CC	and therapy. The invention is useful for treating neuromuscular
CC	disorders, autonomic or central nervous system disorders such as
CC	Parkinson's disease, Huntingtons disease, Alzheimers disease,
CC	schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is
CC	also useful in gene therapy. The present sequence is mouse high affinity
CC	choline transporter (mCHT) cDNA. Note: The sequence data for this patent
CC	did not form part of the printed specification, but was obtained in
CC	electronic format directly from the US patent office at
CC	seqdata.neptco.gov/sequence.html?DocID=US2040248838.
XX	
SEQ	Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;
	Query Match      100.0%; Score 1743; DB 14; Length 1743;
	Best Local Similarity    100.0%; Pred. No. 0;
	Matches 1743; Conservative    0; Mismatches    0; Indels    0; Gaps    0
OY	1    ATGCCCTTCGATGTGAAGAGACTGGTAGCTATTATCCTTTCTACCTCCTTAATTTCTG 60
Db	1    ATGCCTTCCATGTGGAAAGACTGTAAGTATAAATCTTCTTACTCCTTAATTTCTCG 60
OY	61   GTTGGAATATGSGGCTGCATGTGAAAACAAACAACGCGCAACCAGAGAAGCGCACTGAA 120
Db	61   GTTGGAATATGSGGCTGCATGTGAAAACAAACAACGCGCAACCAGAGAAGCGCACTGAA 120
OY	121   GCCATCATAGTCGGGGGGCGGTGACATTGGTTGGTGGTGGTGGTTTACATGACAGCC 180
Db	121   GCATCATAGTCGGGGGGCGGTGACATTGGTTGGTGGTGGTGGTTTACATGACAGCC 180
OY	181   ACCTGGGTTGGAGAGGCTACATCATGTGGAGCAGAGAAAGCAGTGTATGGGCGAGGTTGT 240
Db	181   ACCGTGGTTGGAGAGGCTACATCATGTGGAGCAGAGAAAGCAGTGTATGGGCGAGGTTGT 240
OY	241   GGTCAGCTTGGGCTCATGACACCCATTGGATATCTCTGAGCTAATTTTAACTGCTCG 300
Db	241   GGTCAGCTTGGGCTCATGACACCCATTGGATATCTCTGAGCTAATTTTAACTGCTCG 300
OY	301   TTTTTTGGAAAACCTATGCTTCCCAAGGAAATGTGACTATGTAGAACCTATTCAAACAG 360
Db	301   TTTTTTGGAAAACCTATGCTTCCCAAGGAAATGTGACTATGTAGAACCTATTCAAACAG 360
OY	361   ATTATGGAAGGCGATGSGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTTC 420
Db	361   ATTATGGAAGGCGATGSGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTTC 420
OY	421   TGGGCTGAGAGAAATTTTCTCTGCAATTAAGGGGCAACATCAGGCTGATCATTTGATGGAAT 480
Db	421   TGGGCTGAGAGAAATTTTCTCTGCAATTAAGGGGCAACATCAGGCTGATCATTTGATGGAAT 480
OY	481   GTGAACATATCGGTATGTCCTGCACTCATTTGCATTCCTTTATACCTTAACTGAGGTGGG 540

```

Db      481  GGAAGATATCGGTCATGCTCTGCACTCATGCGCATCTTTATACCCAGTGGGTGGG 540
Qy      541  CTTCTACTCTGTGGCATATCTGATGTTGTTCAGACTATTTCTGCATTTTATAGACCTGTGG 600
Db      541  CTTCTACTCTGTGGCATATCTGATGTTGTTCAGACTATTTCTGCATTTTATAGACCTGTGG 600
Qy      601  ATCAGAGTCCCTTTTGGCCCTGTCACATCTGAGAGTACCGGACATCGGATTCACAGCTGTG 660
Db      601  ATCAGAGTCCCTTTTGGCCCTGTCACATCTGAGAGTACCGGACATCGGATTCACAGCTGTG 660
Qy      661  CATGCTAAATACAGAGTCCCTGCTGGAGACCATTTGAATCAGTTGAAGTCTACACCTGTG 720
Db      661  CATGCTAAATACAGAGTCCCTGCTGGAGACCATTTGAATCAGTTGAAGTCTACACCTGTG 720
Qy      721  CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAAATCCCATGGCAAGCTTCCAGAGG 780
Db      721  CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAAATCCCATGGCAAGCTTCCAGAGG 780
Qy      781  GTCCCTCTTTCATCCCTCAGCAGACCTATGCTCAGGTACTGTCCTTCCGCGAGCTTTGGG 840
Db      781  GTCCCTCTTTCATCCCTCAGCAGACCTATGCTCAGGTACTGTCCTTCCGCGAGCTTTGGG 840
Qy      841  TGCCTGTGATGAGCTCTACCCGCGCATATGAGCTATTTGAGCTTCCACAGACTGG 900
Db      841  TGCCTGTGATGAGCTCTACCCGCGCATATGAGCTATTTGAGCTTCCACAGACTGG 900
Qy      901  AACCAAGATGCTTACAGGGGTATCCAGATCCCAAGACTTAAGAGAGACAGACATGATTTCTC 960
Db      901  AACCAAGATGCTTACAGGGGTATCCAGATCCCAAGACTTAAGAGAGACAGACATGATTTCTC 960
Qy      961  CCGATGCTTCTGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db      961  CCGATGCTTCTGAGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy      1021  TCAGCTGCTGTCATGCTCTCAGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db      1021  TCAGCTGCTGTCATGCTCTCAGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy      1081  CGGAATATCTACAGCTTTCCTTCAAGCAAAATGCAATGCAAGCAAGAAATGTGTGGGTC 1140
Db      1081  CGGAATATCTACAGCTTTCCTTCAAGCAAAATGCAATGCAAGCAAGAAATGTGTGGGTC 1140
Qy      1141  ATGAGGATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db      1141  ATGAGGATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy      1201  ACTGTGTATGAGGCTGTGATCTGATCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db      1201  ACTGTGTATGAGGCTGTGATCTGATCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy      1261  CTGCTCTGTGATCTCTTCAATCAAAAGAACCACTTAATGAGGAGCTGCTGCTGCTGCTGCT 1320
Db      1261  CTGCTCTGTGATCTCTTCAATCAAAAGAACCACTTAATGAGGAGCTGCTGCTGCTGCTGCT 1320
Qy      1321  TTGGAATATCTGGAATATCTGGAAGAGAGCCATATCTTAATCTTGACCCCTTAATC 1380
Db      1321  TTGGAATATCTGGAATATCTGGAAGAGAGCCATATCTTAATCTTGACCCCTTAATC 1380
Qy      1381  TTCTACCCGCTTATTTACTCTGAAGAAATGATATACATCAGAGGTTCCCATTTTAA 1440
Db      1381  TTCTACCCGCTTATTTACTCTGAAGAAATGATATACATCAGAGGTTCCCATTTTAA 1440
Qy      1441  ACTCTCTCATGAGTATCTCATCTTACCAATTTGCTTCTTATCTTATAGCAAGTAT 1500
Db      1441  ACTCTCTCATGAGTATCTCATCTTACCAATTTGCTTCTTATCTTATAGCAAGTAT 1500
Qy      1501  CTAATTTGAAGTGAAGCTTGTCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAGG 1560
Db      1501  CTAATTTGAAGTGAAGCTTGTCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAGG 1560
Qy      1561  CACAGTGAAGAGAACATGAGCAAGACCATTTCTGTGCAAAATGAAATATCAAAATTAAT 1620

```

```

Db      1561  CACAGTGAAGAGAACATGAGCAAGACCATTTCTAGTCAGAAATGAAATATCAAAATTAAT 1620
Qy      1621  GAACCTTGACCTGTGAAACCTGGGAGAGGCTTAACCTTCAGTTCAACTTACCAATTAAG 1680
Db      1621  GAACCTTGACCTGTGAAACCTGGGAGAGGCTTAACCTTCAGTTCAACTTACCAATTAAG 1680
Qy      1681  GAGGCCCTCTCTGATGTTGATTCAGTCCGAGAGGGGTCTGGGACTGAGATAATTTACA 1740
Db      1681  GAGGCCCTCTCTGATGTTGATTCAGTCCGAGAGGGGTCTGGGACTGAGATAATTTACA 1740
Qy      1741  TGA 1743
Db      1741  TGA 1743

RESULT 4
ADV77918
ID ADV77918 standard; cDNA; 1743 BP.
XX
AC ADV77918;
XX
DT 24-FEB-2005 (first entry)
XX
DE Mouse high affinity choline transporter (mCHT) cDNA.
XX
KW Choline transporter; neuromuscular disorder;
KW autonomic nervous system disorder; central nervous system disorder;
KW parkinsons disease; huntingtons chorea; genetic disorder;
KW alzheimers disease; degeneration; neurological disease; schizophrenia;
KW psychiatric disease; myasthenia gravis; immune disorder; gene therapy;
KW cns-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW muscular-gen.; neuroleptic; DNA purification; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..1743
FT FT /*tag= a
FT FT /product= "Mouse CHT protein"
XX
PN US2004248838-A1.
XX
PD 09-DEC-2004.
XX
PF 01-DEC-2003; 2003US-00724806.
XX
PR 23-JUL-2001; 2001US-00911077.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Blakely RD, Apparsundaram S, Ferguson S;
XX
DR MPI; 2005-020586/02.
XX
DR P-PSDB; ADV77919.
XX
PT New human high affinity choline transporter (CHT) cDNA, useful for
PT treating a neuromuscular, autonomic or central nervous system disorder,
PT including Parkinson's disease, Huntington's disease, Alzheimer's,
PT schizophrenia.
XX
PS Example 4; SEQ ID NO 23; 73bp; English.
XX
CC The present invention provides polynucleotides encoding novel high
CC affinity choline transporters (CHTs), methods for their use in screening
CC and therapy. The invention is useful for treating neuromuscular
CC disorders, autonomic or central nervous system disorders such as
CC parkinsons disease, huntingtons disease, alzheimers disease,
CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is
CC also useful in gene therapy. The present sequence is mouse high affinity
CC choline transporter (mCHT) cDNA. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docid=US20040248838.

```

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1743; DB 14; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCTTTCCATGTGGAAGAGCTGGTAGCTATTATTCCTTCTTACCTCTTATTTCTG 60  
DB 1 ATGCTTTCCATGTGGAAGAGCTGGTAGCTATTATTCCTTCTTACCTCTTATTTCTG 60  
OY 61 GTTGGATATGGGCTGCATGAGAAAACAAAACAGCGGCAACCCAGAGAGCGCAGTGA 120  
DB 61 GTTGGATATGGGCTGCATGAGAAAACAAAACAGCGGCAACCCAGAGAGCGCAGTGA 120  
OY 121 GCCATCATAGTCGGGGCCGTGACATTTGGTTTGGTTGGTTTGAACATGACAGCC 180  
DB 121 GCCATCATAGTCGGGGCCGTGACATTTGGTTTGGTTGGTTTGAACATGACAGCC 180  
OY 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAAGACGTGATGGCCAGGTTGT 240  
DB 181 ACCTGGGTTGGAGGAGGCTACATCAATGGGACAGCAAGACGTGATGGCCAGGTTGT 240  
OY 241 GGTCTAGCTGGGGCTCATGSCAACCAATTTGATTTCTGAGTCTAATTTTAGGTCTG 300  
DB 241 GGTCTAGCTGGGGCTCATGSCAACCAATTTGATTTCTGAGTCTAATTTTAGGTCTG 300  
OY 301 TTTTTCGGAAAACCTATGCGTTCGAAGGGATATGTACTATGTAGACCATTCGAACAG 360  
DB 301 TTTTTCGGAAAACCTATGCGTTCGAAGGGATATGTACTATGTAGACCATTCGAACAG 360  
OY 361 ATCTATGGAAGCGATGGGTGGGCTCTTTCATCTCTGCACTGATGGGAGAGATGTTT 420  
DB 361 ATCTATGGAAGCGATGGGTGGGCTCTTTCATCTCTGCACTGATGGGAGAGATGTTT 420  
OY 421 TGGGCTGAGCAATTTTCTGCAATTTAGGGGCCACCATGAGGCTGATCATGATGAT 480  
DB 421 TGGGCTGAGCAATTTTCTGCAATTTAGGGGCCACCATGAGGCTGATCATGATGAT 480  
OY 481 GTGAACATATCGGATCTGCTCTGCACTCATTTGCCATTTCTTATTAACCTTAGTGGTGG 540  
DB 481 GTGAACATATCGGATCTGCTCTGCACTCATTTGCCATTTCTTATTAACCTTAGTGGTGG 540  
OY 541 CTCTACTCTGAGGATATCTGATGTGTCCAGCTATTTGCAATTTTATAGACTGTGG 600  
DB 541 CTCTACTCTGAGGATATCTGATGTGTCCAGCTATTTGCAATTTTATAGACTGTGG 600  
OY 601 ATCAGTGTCCCTTTGGCCCTGTCACATCTGTGAGTCAACGACATGCGATTCACAGCTGTG 660  
DB 601 ATCAGTGTCCCTTTGGCCCTGTCACATCTGTGAGTCAACGACATGCGATTCACAGCTGTG 660  
OY 661 CATGCTAAATACCAAGTCCCTGTGGCTGGAGAACCATTTGAATCAAGTTACACTGG 720  
DB 661 CATGCTAAATACCAAGTCCCTGTGGCTGGAGAACCATTTGAATCAAGTTACACTGG 720  
OY 721 CTGATTAATTTCTGTATTTGATGTGGGTGGAATCCCATGGCAACCTTCCAGAGG 780  
DB 721 CTGATTAATTTCTGTATTTGATGTGGGTGGAATCCCATGGCAACCTTCCAGAGG 780  
OY 781 GTCTCTCTTCATCTCCAGCCACCTATGCTCAGGTACTCTCTTCTGACAGCTTTTGGG 840  
DB 781 GTCTCTCTTCATCTCCAGCCACCTATGCTCAGGTACTCTCTTCTGACAGCTTTTGGG 840  
OY 841 TGCCGTGTGATGCTCTACCCGCAATATGCAATAGAGAGCTTATGGAGCTTCCACAGCTGG 900  
DB 841 TGCCGTGTGATGCTCTACCCGCAATATGCAATAGAGAGCTTATGGAGCTTCCACAGCTGG 900  
OY 901 AACCAAGACTGCTTAAGGGGATCCAGATCCCAAGACTTAAGAGGAAGCAACATGATTTCTC 960  
DB 901 AACCAAGACTGCTTAAGGGGATCCAGATCCCAAGACTTAAGAGGAAGCAACATGATTTCTC 960  
OY 961 CCGATGCTTCTGACATGCTGACCTGCTGTGATCATCTCTTCTTTGGGCTGTGCTGTT 1020  
DB 961 CCGATGCTTCTGACATGCTGACCTGCTGTGATCATCTCTTCTTTGGGCTGTGCTGTT 1020

DB 961 CCGATGCTTCTGACATGCTGACCTGCTGTGATCATCTCTTCTTTGGGCTGTGCTGTT 1020  
OY 1021 TCAGCTGTGATCATGCTCTCAGCTGACTGTCATCTGTCATCTGTCGCGAGTTCTATGTTTCT 1080  
DB 1021 TCAGCTGTGATCATGCTCTCAGCTGACTGTCATCTGTCATCTGTCGCGAGTTCTATGTTTCT 1080  
OY 1081 CGGAATATCTACAGCTTCTCTTCAAGAAAATGATACAGCAAGAAATGTGTGGGTC 1140  
DB 1081 CGGAATATCTACAGCTTCTCTTCAAGAAAATGATACAGCAAGAAATGTGTGGGTC 1140  
OY 1141 ATGAGATACACTGTGCTTGTGTGAGACATCTGCAACAGCCATGCTTTGCTGACGAAG 1200  
DB 1141 ATGAGATACACTGTGCTTGTGTGAGACATCTGCAACAGCCATGCTTTGCTGACGAAG 1200  
OY 1201 ACTGTGATGAGGCTCTGCTGATCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260  
DB 1201 ACTGTGATGAGGCTCTGCTGATCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260  
OY 1261 CTGCTCTGTGATCTCTTATCAAGAGAACCAACATTAATGGGGCAGTTGCTGTATATTT 1320  
DB 1261 CTGCTCTGTGATCTCTTATCAAGAGAACCAACATTAATGGGGCAGTTGCTGTATATTT 1320  
OY 1321 TTTGACTATTCTCTGAGATTAATCTGAGAGAGCCATCTATTAATCTGACGCCCTTAATC 1380  
DB 1321 TTTGACTATTCTCTGAGATTAATCTGAGAGAGCCATCTATTAATCTGACGCCCTTAATC 1380  
OY 1381 TTTACCCCTGTTATTACTCTGACAAAGATGTATATCAATCAGAGGTTCCCATTTTAA 1440  
DB 1381 TTTACCCCTGTTATTACTCTGACAAAGATGTATATCAATCAGAGGTTCCCATTTTAA 1440  
OY 1441 ACTCTCTCATAGTGTAACTCATTTCTTACCAAACTTGTGTTCTTATCTGACCGCAAGTAT 1500  
DB 1441 ACTCTCTCATAGTGTAACTCATTTCTTACCAAACTTGTGTTCTTATCTGACCGCAAGTAT 1500  
OY 1501 CTATTGGAAGTGAACCTTGCCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAG 1560  
DB 1501 CTATTGGAAGTGAACCTTGCCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAG 1560  
OY 1561 CACAGTGAAGAGACATGAGACCAATCTTCAATTTAGATGATTTGATGCTGTGTGCGAAG 1620  
DB 1561 CACAGTGAAGAGACATGAGACCAATCTTCAATTTAGATGATTTGATGCTGTGTGCGAAG 1620  
OY 1621 GAACCTTGACCTGTGAACCTGCGAGAGCCTTAACCTCAAGTCAACTTTACCAAAATAG 1680  
DB 1621 GAACCTTGACCTGTGAACCTGCGAGAGCCTTAACCTCAAGTCAACTTTACCAAAATAG 1680  
OY 1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGGGGCTGAGGACTGAAGATTAATTACAA 1740  
DB 1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGGGGCTGAGGACTGAAGATTAATTACAA 1740  
OY 1741 TGA 1743  
DB 1741 TGA 1743

RESULT 5  
AAFB1713  
ID AAFB1713 standard; cDNA; 1743 BP.  
XX  
AC AAFB1713;  
XX  
DT 01-JUN-2001 (first entry)  
XX  
DE Mouse high affinity choline transporter protein encoding cDNA.  
XX  
KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;  
XX  
OS Mus musculus.  
XX  
FH Key 1.1743 Location/Qualifiers  
FT CDS  
FT tag= a

FT /product= "high affinity choline transporter"  
XX MO200116315-A1.  
XX 08-MAR-2001.  
XX 18-AUG-2000; 2000WO-JP005545.  
XX 27-AUG-1999; 99JP-00240642.  
XX 27-DEC-1999; 99JP-00368991.  
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX Haga T, Okuda T;  
XX WPI: 2001-326688/23.  
XX P-PSDB; AAB74666.  
XX  
XX New rat and human spinal cord high affinity choline transporters, useful  
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for  
PT treating Alzheimer's disease.  
XX  
XX Claim 12; Page 78-82; 90pp; Japanese.  
XX  
XX The present sequence encodes a mouse (Mus musculus) high affinity choline  
CC transporter protein designated cho-1. The cho-1 protein has nootropic and  
CC neuroprotective activities. The cho-1 polynucleotide and protein can be  
CC used for the diagnosis of diseases related to the expression of cho-1 by  
CC comparing the cho-1 polynucleotide sequence in a sample to that of a  
CC control. Drug compositions containing the cho-1 protein or expression  
CC promoters or inhibitors of cho-1 are useful for treating disorders  
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease  
XX  
SQ Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;  
  
Query Match 99.3%; Score 1730.2; DB 4; Length 1743;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 ATGCCCTTCCATGTGGAAGAGCTGTGATATATCTCTTCACTCTCTTATATTTCTG 60  
DB 1 ATGCTCTTCCAGTGAAGAGCTGTGATATATCTCTTCACTCTCTTATATTTCTG 60  
QY 61 GTTGAATATGGGCTCATGAAAAACCAAAAACAGGGCAACCCAGAAAGCGCAGTGA 120  
DB 61 GTTGAATATGGGCTCATGAAAAACCAAAAACAGGGCAACCCAGAAAGCGCAGTGA 120  
QY 121 GCCATCATAGTCGGGGGCGGTGACATTTGTTGTTGGTGTGTTTACCATGACAGCC 180  
DB 121 GCCATCATAGTCGGGGGCGGTGACATTTGTTGTTGGTGTGTTTACCATGACAGCC 180  
QY 181 ACCTGGGTGGAGAGGCTACATCAATGGGACAGCAAGCACTGTATGGGCGAGTTGT 240  
DB 181 ACCTGGGTGGAGAGGCTACATCAATGGGACAGCAAGCACTGTATGGGCGAGTTGT 240  
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTG 300  
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTG 300  
QY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACATGTTTAGACCATTTCAAAAG 360  
DB 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACATGTTTAGACCATTTCAAAAG 360  
QY 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCATCCCTGCACTGATGGAGAGATGTC 420  
DB 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCATCCCTGCACTGATGGAGAGATGTC 420  
QY 421 TGGGCTGACGCAATTTTCTCTGCAATTAAGGGCCACCATGACGCTGATCAATGTGAT 480  
DB 421 TGGGCTGACGCAATTTTCTCTGCAATTAAGGGCCACCATGACGCTGATCAATGTGAT 480  
QY 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTAGTGGGTGG 540  
DB 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTAGTGGGTGG 540

DB 481 GTGAACATATCGGTCTTGTCTTGCACTCATTTGCACTTTTATACCTAGTGGGTGG 540  
QY 541 CTCTACTCTGTCGCAATATATGATGTGTCAGGCTATCTGATTTTATAGACGTGG 600  
DB 541 CTCTACTCTGTCGCAATATATGATGTGTCAGGCTATCTGATTTTATAGACGTGG 600  
QY 601 ATCAGTGTCCCTTTTGGCCCTGTACATCCTGACAGTCAACCGATCGGATTCACAGCTGTG 660  
DB 601 ATCAGTGTCCCTTTTGGCCCTGTACATCCTGACAGTCAACCGATCGGATTCACAGCTGTG 660  
QY 661 CATGCTAAATACAGAGTCCCTGGGCAACATTGAATCACTGTAAGTCTACACTGG 720  
DB 661 CATGCTAAATACAGAGTCCCTGGGCAACATTGAATCACTGTAAGTCTACACTGG 720  
QY 721 CTTGATTAATTTTCTGTTATGTATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAG 780  
DB 721 CTTGATTAATTTTCTGTTATGTATGCTGGGTGGAATCCCATGGCAAGCCTACTTCCAGAG 780  
QY 781 GTCTCTCTTTCATCTCTACAGCCACTATGCTCAGTACTGTCTTCTGACAGCTTTGGG 840  
DB 781 GTCTCTCTTTCATCTCTACAGCCACTATGCTCAGTACTGTCTTCTGACAGCTTTGGG 840  
QY 841 TGCCTGTGATGAGCTCTTACCCGCAATATGCAATAGAGCTATTGGAGCTTCCACAGACTGG 900  
DB 841 TGCCTGTGATGAGCTCTTACCCGCAATATGCAATAGAGAGCTATTGGAGCTTCCACAGACTGG 900  
QY 901 AACCGACATGCTTACGGGATACCAATCCCAAGACTTAAGAGAGGACAGACATGATTTCTC 960  
DB 901 AACCGACATGCTTACGGGATACCAATCCCAAGACTTAAGAGAGGACAGACATGATTTCTC 960  
QY 961 CCGATGCTTCTCAGTACTCTGCTGCTGTATCATCTCTTGGGCTTGGTGTCTGT 1020  
DB 961 CCGATGCTTCTCAGTACTCTGCTGCTGTATCATCTCTTGGGCTTGGTGTCTGT 1020  
QY 1021 TTAGCTGTGATGATGCTCTCACTGATCTGTCATCTGTGCGGAGTTCTATGTTTGTCT 1080  
DB 1021 TTAGCTGTGATGATGCTCTCACTGATCTGTCATCTGTGCGGAGTTCTATGTTTGTCT 1080  
QY 1081 CGGAATATCTACAGGTTTCCCTTCAGCAAAATGCAATGCAAGAAATTTGTTGGGTCT 1140  
DB 1081 CGGAATATCTACAGGTTTCCCTTCAGCAAAATGCAATGCAAGAAATTTGTTGGGTCT 1140  
QY 1141 ATGAGGATCATCTGTGTGTGTTGGAGCATCTGCAACGCCATGGCTTTGTGCAAGAG 1200  
DB 1141 ATGAGGATCATCTGTGTGTGTTGGAGCATCTGCAACGCCATGGCTTTGTGCAAGAG 1200  
QY 1201 ACTGTGTATGGGCTCTGTTACTGAGCTGTACCTTTGTCTAATCATCTTCTCCACAG 1260  
DB 1201 ACTGTGTATGGGCTCTGTTACTGAGCTGTACCTTTGTCTAATCATCTTCTCCACAG 1260  
QY 1261 CTGCTCTGTGTACTCTCTCAACAAAGAACCAACATGATGGGGCAGTTGCTGTTATTT 1320  
DB 1261 CTGCTCTGTGTACTCTCTCAACAAAGAACCAACATGATGGGGCAGTTGCTGTTATTT 1320  
QY 1321 TTTGACATATCTGAGAAATTACTGGAGAGAGCCATATCTATCTTGGACGCTTAAATC 1380  
DB 1321 TTTGACATATCTGAGAAATTACTGGAGAGAGCCATATCTATCTTGGACGCTTAAATC 1380  
QY 1381 TTCTACCTGTTTATCTCTGACAAAGATGATATACATCAAGAGTTCCCATTTAA 1440  
DB 1381 TTCTACCTGTTTATCTCTGACAAAGATGATATACATCAAGAGTTCCCATTTAA 1440  
QY 1441 ACTCTCTCCAGGTTAATCTTCTTACCAACATTTGTTTCTTATCTAGCAAGATAT 1500  
DB 1441 ACTCTCTCCAGGTTAATCTTCTTACCAACATTTGTTTCTTATCTAGCAAGATAT 1500  
QY 1501 CTATTTGAAAGTGAACTTGGCTTCCAAATATGATATTTGATGCTGTGTGCAAG 1560  
DB 1501 CTATTTGAAAGTGAACTTGGCTTCCAAATATGATATTTGATGCTGTGTGCAAG 1560  
QY 1561 CACAGTGAAGAGAAATGAGAACATGCAAGCACTTCAAGTCAAGAAATTCAAATTAAT 1620  
DB 1561 CACAGTGAAGAGAAATGAGAACATGCAAGCACTTCAAGTCAAGAAATTCAAATTAAT 1620

```
QY 1621 GAACCTGACCTGTGAAACCTTCGGCAGAGCCTTAACCTCAGTTCACTTACCATTAAG 1680
    |||
Db 1621 GAACCTGACCTGTGAAACCTTCGGCAGAGCCTTAACCTCAGTTCACTTACCATTAAG 1680
QY 1681 GAGGCCCCCTTGTGATGATTCAGTCCGGAGGGGCTTCGGACTGGAAGATTAATTACAA 1740
    |||
Db 1681 GAGGCCCCCTTGTGATGATTCAGTCCGGAGGGGCTTCGGACTGGAAGATTAATTACAA 1740
QY 1741 TGA 1743
    |||
Db 1741 TGA 1743

RESULT 6
AAD02457
ID AAD02457 standard; cDNA; 4938 BP.
AC AAD02457;
XX
XX 24-APR-2001 (first entry)
DT
XX
XX Mouse P4P6B1 OMA (obese mice adipocyte) protein encoding cDNA.
DE
XX
XX Mouse; OMA protein; obese mice adipocyte; P4P6B1;
KW fuel metabolism disorder; therapy; obesity; diabetes; gene therapy;
KM anorectic; antidiabetic; ss.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH 247..1989
FT CDS /*tag= a
FT misc_feature /product= "Mouse P4P6B1 OMA protein"
FT 988..1342
FT /*tag= b
FT /note= "Portion of original 450 bp PCR fragment"
XX
XX WO200078950-A2.
XX
XX 28-DEC-2000.
XX
XX 13-JUN-2000; 2000WO-US016217.
XX
XX 22-JUN-1999; 99US-0141515P.
XX
XX (AMYL-) AMYLIN PHARM INC.
PA
PI Sierzege M, Albrandt K;
PI WPI; 2001-112322/12.
DR P-PDB; AA172388.
XX
XX Novel obese mice adipocyte polypeptides useful in diagnosis and treatment
PT of disorders of fuel metabolism such as obesity or diabetes.
XX
XX Claim 2; Fig 3; 83pp; English.
XX
XX The present sequence is mouse P4P6B1 cDNA which encodes OMA (obese mice
CC adipocyte) protein. The P4P6B1 fragment was generated by RNA
CC fingerprinting using random primers P4 and P6. OMA is used as a
CC diagnostic reagent for diagnosing a disorder of fuel metabolism in an
CC underweight or an overweight individual, by detecting the transcription
CC level of a gene encoding OMA, which is induced or repressed in an
CC individual by a factor such as genetic obesity, fasting and refeeding of
CC a fasted individual. OMA is useful in the generation of antibodies, for
CC use in pharmaceutical compositions and for studying DNA/protein
CC interactions. Nucleic acids encoding OMA are involved in gene therapy. An
CC inhibitor of OMA or an antisense oligonucleotide that inhibits expression
CC of OMA are useful for treating disorders of fuel metabolism such as
CC obesity or diabetes
XX
XX Sequence 4938 BP; 1436 A; 1012 C; 976 G; 1514 T; 0 U; 0 Other;
```

```
Query Match 99.3%; Score 1730.2; DB 5; Length 4938;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1735; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGCGCTTCCATGAGGAGGACGTGTAGCTATTAATCCCTTCTCACTTCCTATATTTCTG 60
    |||
Db 247 ATGCTTTCCACGTAGAGAGACGTGTAGCTATTAATCCCTTCTCACTTCCTATATTTCTG 306
QY 61 GTTGAATATATGGGCTGCATGGAAAAACCAAAACAGCGGCAACCAAGAGCGCAGTGA 120
    |||
Db 307 GTTGAATATATGGGCTGCATGGAAAAACCAAAACAGCGGCAACCAAGAGCGCAGTGA 366
QY 121 GCCATCATAGTCGGGGGCGCGTGAATCTGTTGTTGTTGTTGTTGTTTACATGACACC 180
    |||
Db 367 GCCATCATAGTCGGGGGCGCGTGAATCTGTTGTTGTTGTTGTTTACATGACACC 426
QY 181 ACCTGGGTTGAGAGAGGCTTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
    |||
Db 427 ACCTGGGTTGAGAGAGGCTTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 241 GGTCTAGCTTGGGCTCATGACACCATGGATATTTCTGAGTCTTAATTTAGTGTCTG 300
    |||
Db 487 GGTCTAGCTTGGGCTCATGACACCATGGATATTTCTGAGTCTTAATTTAGTGTCTG 546
QY 301 TTTTGTGGAACCTATGCGTTCCAGAGGATATGTAATGTAATGTAATGTAATGTAATG 360
    |||
Db 547 TTTTGTGGAACCTATGCGTTCCAGAGGATATGTAATGTAATGTAATGTAATGTAATG 606
QY 361 ATCTATGGAAGAGCGATGGGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGTC 420
    |||
Db 607 ATCTATGGAAGAGCGATGGGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGTC 666
QY 421 TGGGCTGAGCAATTTTCTGCAATGAGGGCCACATCAGCGTATATTGATGTGAT 480
    |||
Db 667 TGGGCTGAGCAATTTTCTGCAATGAGGGCCACATCAGCGTATATTGATGTGAT 726
QY 481 GTGAACATATGGGTATGTCCTGCACTCATGTCATCTTTTATCCCTGTGGGTGGG 540
    |||
Db 727 GTGAACATATGGGTATGTCCTGCACTCATGTCATCTTTTATCCCTGTGGGTGGG 786
QY 541 CTCTACTGTGAGCATATAGATGATGTCAGATATCTGACATTTTATAGACTGTGG 600
    |||
Db 787 CTCTACTGTGAGCATATAGATGATGTCAGATATCTGACATTTTATAGACTGTGG 846
QY 601 ATCAGTGTCCCTTTTGGCTGTCAATCTGCAATCTGCAATCTGCAATCTGCAATCTG 660
    |||
Db 847 ATCAGTGTCCCTTTTGGCTGTCAATCTGCAATCTGCAATCTGCAATCTGCAATCTG 906
QY 661 CATGCTAATATACAGAGTCCCTGGGCTGGGAGCAATGTAATGTAATGTAATGTAATG 720
    |||
Db 907 CATGCTAATATACAGAGTCCCTGGGCTGGGAGCAATGTAATGTAATGTAATGTAATG 966
QY 721 CTGTATATATTTCTGTATTAATGATGTGGGTGAAATCCCATGGCAAGCCTTCTCAGAG 780
    |||
Db 967 CTGTATATATTTCTGTATTAATGATGTGGGTGAAATCCCATGGCAAGCCTTCTCAGAG 1026
QY 781 GTCTCTCTTCACTCTACGCCACCTATGCTCAGGTACTGCTCTTCTGCACTTTTGGG 840
    |||
Db 1027 GTCTCTCTTCACTCTACGCCACCTATGCTCAGGTACTGCTCTTCTGCACTTTTGGG 1086
QY 841 TGCCGTGATAGGGCTCTACCGGCCATATGATGATGAGATTTGGAGCTTCCACAGACTGG 900
    |||
Db 1087 TGCCGTGATAGGGCTCTACCGGCCATATGATGATGAGATTTGGAGCTTCCACAGACTGG 1146
QY 901 AACCAAGACTGCTACGGGTATCCAGATCCAAAGCTAAGAGGAGAGAGAGAGATTTCTC 960
    |||
Db 1147 AACCAAGACTGCTACGGGTATCCAGATCCAAAGCTAAGAGGAGAGAGAGAGATTTCTC 1206
QY 961 CCGATGCTTTCGCACTACTCTGCGCTGTGATCATCTCTTCTTGGGCTGTGCTGT 1020
    |||
Db 1207 CCGATGCTTTCGCACTACTCTGCGCTGTGATCATCTCTTCTTGGGCTGTGCTGT 1266
```

QY	1021	TCAGCTGCTGCATGTCCTCTCAGAGCTACCTGCTACATCCCTGTCGGGAGATTATGTTTCT	1080
Dp	1267	TCAGTGTCTGCATGTCCTCTCAGCTACCTGCTACATCCCTGTCGGGAGATTATGTTTCT	1328
QY	1081	CGGAATATCTACCAAGCTTTCCTTCAGACAAATGCAATGCAACAAAGAAATGTGTGGCT	1140
Dp	1327	CGGAATATCTACCAAGCTTTCCTTCAGACAAATGCAATGCAACAAAGAAATGTGTGGCT	1386
QY	1141	ATGAGATCATCTGTCCTGTGTGTGGAGACATCTGCACAGCCATGCTTTGTCTGACGAG	1200
Dp	1387	ATGAGATCATCTGTCCTGTGTGTGGAGACATCTGCACAGCCATGCTTTGTCTGACGAG	1448
QY	1201	ACTGNGATGAGGCTCTGTGCTACCTGAGCTCTGAGCTTGTCTACATCATCATCTTCCACAG	1266
Dp	1447	ACTGNGATGAGGCTCTGTGCTACCTGAGCTCTGAGCTTGTCTACATCATCATCTTCCACAG	1506
QY	1261	CTGCTCTGTGTACTCTTTCATCAAAAGGAAACCAACATTATGAGGAGCTGTGTATATTT	1320
Dp	1507	CTGCTCTGTGTACTCTTTCATCAAAAGGAAACCAACATTATGAGGAGCTGTGTATATTT	1566
QY	1321	TTTGGACATATTCTCTGAGATTTACTGAGAGAGACCATTTCTATACTTCCAGCCCTTAATC	1380
Dp	1567	TTTGGACATATTCTCTGAGATTTACTGAGAGAGACCATTTCTATACTTCCAGCCCTTAATC	1628
QY	1381	TTTCAACCTGGTTATTAATCTCTGACAAAGATGTATATCATTCAGAGTTCCCATTTTAA	1440
Dp	1627	TTTCAACCTGGTTATTAATCTCTGACAAAGATGTATATCATTCAGAGTTCCCATTTTAA	1686
QY	1441	ACTCTCTCCATGTTACCTCAATTTCTTTACCAACATTTGTGTCTTTATCTACGCAAGTAT	1500
Dp	1687	ACTCTCTCCATGTTACCTCAATTTCTTTACCAACATTTGTGTCTTTATCTACGCAAGTAT	1746
QY	1501	CTATTTTGAAAGTGAACCTTGTGCTTCGCTTCGAAATTTAAGTATTTGATGCTGTGTGCGAAG	1560
Dp	1747	CTATTTTGAAAGTGAACCTTGTGCTTCGCTTCGAAATTTAAGTATTTGATGCTGTGTGCGAAG	1806
QY	1561	CACAGTGAAGGAACAATGAGCAAGACATTTCTAGTCACAAAATGAAATATCAAAATTTAAT	1620
Dp	1807	CACAGTGAAGGAACAATGAGCAAGACATTTCTAGTCACAAAATGAAATATCAAAATTTAAT	1866
QY	1621	GAACTTGCACCTGTGAAACCTGTGGAGAGCCTTAACCTCAGTTCACATTTGACCAATTAAG	1680
Dp	1867	GAACTTGCACCTGTGAAACCTGTGGAGAGCCTTAACCTCAGTTCACATTTGACCAATTAAG	1926
QY	1681	GAGGCTCTCTTTGATGTTGATTCCAGTTCGGAGGGGCTGTGGACTGAAGATTAATTACAA	1740
Dp	1927	GAGGCTCTCTCTTTGATGTTGATTCCAGTTCGGAGGGGCTGTGGACTGAAGATTAATTACAA	1986
QY	1741	TGA 1743	
Dp	1987	TGA 1989	

PN	WC0200116315-AL.
PM	
PD	
PC	08-MAR-2001.
PF	
PE	18-AUG-2000; 2000WO-JP005545.
XX	
XX	27-AUG-1999; 99JP-00240642.
PR	27-DEC-1999; 99JP-00368991.
XX	
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX	
P1	Haga T, Okuda T;
XX	
XX	WPI: 2001-226688/23.
DR	P-PSDB; AAB74664.
XX	
PT	New rat and human spinal cord high affinity choline transporters, useful
PT	in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT	treating Alzheimer's disease.
XX	
XX	Claim 6; Page 64-68; 90pp; Japanese.
XX	
CC	The present sequence encodes a rat ( <i>Rattus norvegicus</i> ) high affinity
CC	choline transporter protein designated cho-1. The cho-1 protein has
CC	neurotropic and neuroprotective activities. The cho-1 polynucleotide and
CC	nucleotide can be used for the diagnosis of diseases related to the
CC	expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
CC	sample to that of a control. Drug compositions containing the cho-1
CC	protein or expression promoters or inhibitors of cho-1 are useful for
CC	treating disorders characterized by abnormal levels of cho-1, such as
CC	Alzheimer's disease
XX	
SO	Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;

Query Match	92.7%;	Score 1616.6;	DB 4;	Length 1743;
Best Local Similarity	95.5%;	Pred. No. 0;		
Matches 1664; Conservative	0;	Mismatches 79;	Indels 0;	Gaps 0;

Qy	1	ATGCTTTCCATGTGGGAAGGAGCTGGTAGTATATATCTCTTACCTCTCTTAATATTCG	60
Db	1	ATGCTTTCCATGTGAAGAGCACTAGTAGATTAATCTGTCTTACCTCTCTTAATATTCG	60
Qy			
Db			
Qy	61	GTGGAAATATGGGCTGCATGTGAAAACCAAAACAGCGCACACCAGAAAGCGCACTGAA	120
Db	61	GTGGAAATATGGGCTGCATGTGAAAACCAAAACAGCGGTAATGCAAAAGAACCCAGGAA	120
Qy	121	GCCATCATAGTCGGGGGGCCGATGACATTTGGTTGTTGGTGTGGTGTATACATGACAGCC	180
Db	121	GCCATCATAGTTGGGGGGCCGAGACATTTGGTTTGTGGTTGGTGTATACATGACAGCC	180
Qy	181	ACCTGGGTTGGAGAGAGGCTATCATCATGGGACAGAGAGAGAGAGAGTGTGT	240
Db	181	ACCTGGGTTGGAGAGAGGTTATCATCAACGGGACAGCTGAAGCAGTTATGGGACAGGTTGT	240
Qy	241	GCTCTAGCTTGGGCTCATGACCAACCATTTGGATTTCTCTGAGTCTAATTTTAGTGGTCTG	300
Db	241	GCTCTAGCTTGGGCTCAGGACCAACCATTTGGATTTCTCTGAGTCTGATTTTAGTGGCTCTG	300
Qy	301	TTTTTTGGGAAACCTATGCGTTCCAAGGGATATGTGACTATGTATGAGACCATTTCAACAG	360
Db	301	TTTTTTGCAAAACCTATGCGTTCCAAAGGATATGTGACTATGTATGAGACCCGTTCAACAG	360
Qy	361	ATCTATGGAAGCGGATGGGTGGGTGCTCTTCAATCCCTGCACCTGATGGGAGAGATCTTC	420
Db	361	ATCTATGGAAGCGGATGGGTGGGTGCTCTTCAATCCCTGCACCTGATGGGAGAGATCTTC	420
Qy	421	TGGGCTGAGCAATTTTCTCTGCATTTAGGGGCCACATCAGCGTATCTTATATGATGAT	480
Db	421	TGGGCTGAGCAATTTTCTCTGCATTTAGGGGGCTACATCAGCGTATCATTATGATGAT	480
Qy	481	GTGAACAATATGGGTATCTCTCTGCACCTCATTTGCACATTTCTTAAATACCTAGTGGGTGG	540
Db	481	GTGAACAATATGGGTATCTCTCTGCACCTCATTTGCACATTTCTTAAATACCTCTGGGGAGGG	540



```

Qy 541 CTCTACTGTGGCATATATCTGATGTGGTCAGCTAATTCGATTTTATATAGAGCTGG 600
Db 541 CTCTACTGTGGCATATATCTGATGTGGTCAGCTAATTCGATTTTATATAGAGCTGG 600
Qy 601 ATCAATGTCCTTTGGCCCTGTACATCTGTGAGTCACCGACATGGGATTCAGCTGG 660
Db 601 ATCAATGTCCTTTGGCCCTGTACATCTGTGAGTCACCGACATGGGATTCAGCTGG 660
Qy 661 CATGCTAATAACAGAGTCCCTGGCTGGGAACCATTTGATCAGTTGAGTACACTGG 720
Db 661 CATGCTAATAACAGAGTCCCTGGCTGGGAACCATTTGATCAGTTGAGTACACTGG 720
Qy 721 CTGTAATATTTCTGTATTATGATGCTGGGTGAATCCAGTCCAGCTTCTTCCAGAG 780
Db 721 CTGTAATATTTCTGTATTATGATGCTGGGTGAATCCAGTCCAGCTTCTTCCAGAG 780
Qy 781 GTCCCTCTCTTCATCTTCAGCCACCTATGCTCAGTACTGTCTTCTGTGGCAGCTTTGGG 840
Db 781 GTCCCTCTCTTCATCTTCAGCCACCTATGCTCAGTACTGTCTTCTGTGGCAGCTTTGGG 840
Qy 841 TGCCCTGTGATGGCTCTACCCGCATATGCTATGAGAGCTTATGAGCTTCCACAGACTGG 900
Db 841 TGCCCTGTGATGGCTCTACCCGCATATGCTATGAGAGCTTATGAGCTTCCACAGACTGG 900
Qy 901 AACCAAGACTGCTACGCGGTATCCAGATCCCAAGACTTAGAGAGAGACAGACATGATTC 960
Db 901 AACCAAGACTGCTACGCGGTATCCAGATCCCAAGACTTAGAGAGAGACAGACATGATTC 960
Qy 961 CGGATGCTTGGCAGTACCTGCGCCCTGTGTACATCTCTTGGTGGGTGGTGTGT 1020
Db 961 CGGATGCTTGGCAGTACCTGCGCCCTGTGTACATCTCTTGGTGGGTGGTGTGT 1020
Qy 1021 TCAGCTGTGTCATGTCCTCAGCTGATGTCGTCATCTGTCGCGAGTCTATGTTGGCT 1080
Db 1021 TCAGCTGTGTCATGTCCTCAGCTGATGTCGTCATCTGTCGCGAGTCTATGTTGGCT 1080
Qy 1081 CGGAATATCTACACAGCTTCTTCCAGCAAATATGATCAGACAGAGAAATGTGTGGTCT 1140
Db 1081 CGGAATATCTACACAGCTTCTTCCAGCAAATATGATCAGACAGAGAAATGTGTGGTCT 1140
Qy 1141 ATGAGAGATCAGTGTGTGTGGAGATCTGCAAGAGCCATGGCTTGTGTCAGAG 1200
Db 1141 ATGAGAGATCAGTGTGTGTGGAGATCTGCAAGAGCCATGGCTTGTGTCAGAG 1200
Qy 1201 ACTGTGTATGGGCTGTGATCTGAGCTGAGCTTGTCTACATCATCTTCCACAG 1260
Db 1201 ACTGTGTATGGGCTGTGATCTGAGCTGAGCTTGTCTACATCATCTTCCACAG 1260
Qy 1261 CTGCTCTGTGTACTCTTCAACAAAGAACCACTTATGGGCGAGTCTGTGTTATAT 1320
Db 1261 CTGCTCTGTGTACTCTTCAACAAAGAACCACTTATGGGCGAGTCTGTGTTATAT 1320
Qy 1321 TTGGACTATTCCTGAGATTACTGGAGAGAGCCATCTATTACTTGGAGCCCTTAATC 1380
Db 1321 TTGGACTATTCCTGAGATTACTGGAGAGAGCCATCTATTACTTGGAGCCCTTAATC 1380
Qy 1381 TTCTACCCCTGTTATTAATCTCTGACAAAGATGATATACATCAGAGTTCCTATTAA 1440
Db 1381 TTCTACCCCTGTTATTAATCTCTGACAAAGATGATATACATCAGAGTTCCTATTAA 1440
Qy 1441 ACTCTCTCGATGTTACCTGATCTTTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500
Db 1441 ACTCTCTCGATGTTACCTGATCTTTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500
Qy 1501 CTATTGGAAGTGAAGCTTGGCTCCAAATATAGATGATTTGATGCTGTGGCAGAG 1560
Db 1501 CTATTGGAAGTGAAGCTTGGCTCCAAATATAGATGATTTGATGCTGTGGCAGAG 1560
Qy 1561 CACAGTGAAGAGACATGAGCAAGACCATTTAGTCAGAAATGAAAACATCAATTAAT 1620
Db 1561 CACAGTGAAGAGACATGAGCAAGACCATTTAGTCAGAAATGAAAACATCAATTAAT 1620

```

```

Qy 1621 GAACCTGACCTGTGAAACCTGGGAGAGCCTTAACCTTCAGTTCACTTTCACAAATAAG 1680
Db 1621 GAACCTGACCTGTGAAACCTGGGAGAGCCTTAACCTTCAGTTCACTTTCACAAATAAG 1680
Qy 1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGAGGGGTGGGACATGAGATTAATTAACA 1740
Db 1681 GAGGCCCTCTTGAATGATTTCCAGTCCGAGAGGGATCTGGGACATGAGATTAATTAACA 1740
Qy 1741 TGA 1743
Db 1741 TGA 1743

RESULT 8
ADD50642
ID ADD50642 standard; cDNA, 4904 BP.
XX
AC ADD50642;
XX
DT 15-JAN-2004 (first entry)
XX
DE cDNA encoding rat high-affinity choline transporter (rCHT).
XX
KW Rat, high-affinity choline transporter; rCHT; cholinergic function;
KW Parkinson's disease; Huntington's disease; Alzheimer's disease;
KW schizophrenia; dyskinesia; myasthenia gravis; brain;
KW cholinergic signaling; antiparkinsonian; anticonvulsant; nootropic;
KW neuroprotective; neuroleptic; gene; ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT CDS 224..1966
FT /tag= a
FT /product= "rCHT"
XX
PN US2003114399-A1.
XX
PD 19-JUN-2003.
XX
PF 23-JUL-2001; 2001US-00911077.
XX
PR 23-JUL-2001; 2001US-00911077.
XX
PA (BLAKELY R D.
PA (APPA/) APPARUNDARAM S.
PA (FERG/) FERGUSON S.
XX
PI Blakely RD, Apparsundaram S, Ferguson S;
XX
DR WPI; 2003-810914/76.
XX
DR P-PSDB; ADD50643.
XX
PT Novel isolated polynucleotide encoding human or mouse high affinity
PT choline transporter polypeptide, useful in gene therapy to increase
PT cholinergic function in a cell of a patient suffering from Alzheimer's
PT disease.
XX
PS Example 1; SEQ ID NO 5; 74pp; English.
XX
CC The present invention relates to the isolation of polynucleotide
CC sequences encoding human and mouse high-affinity choline transporter
CC (hCHT and mCHT respectively), and the proteins they encode. The gene
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT
CC polynucleotide sequence when delivered to a cell, increases cholinergic
CC function in the cell that is in a patient having Parkinson's disease,
CC Huntington's disease, Alzheimer's disease, schizophrenia, dyskinesia or
CC myasthenia gravis. The hCHT antibody is useful for controlling
CC transporter CHT proteins in the brain, and for treating the above
CC mentioned diseases. The antibody is also useful for diagnosing the above
CC mentioned disorders and to detect the influence of cholinergic
CC signalling. The present sequence encodes rat CHT (rCHT). Note: The

```



CC sequence data for this patent was obtained in electronic format directly  
 from the USPTO web site at [segdata.uspto.gov](http://segdata.uspto.gov).

XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

Query Match 92.7%; Score 1616.6; DB 10; Length 4904;  
 Best Local Similarity 95.5%; Pred. No. 0;  
 Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

QY 1 ATGCTTTTCATGTGGAAGAGCTGAGTATTTCTCTTCTACTCTCTTATTTCTG 60
DB ATGCTTTTCATGTGGAAGAGCTGAGTATTTCTCTTCTACTCTCTTATTTCTG 283
QY 61 GTTGAATATGAGGCTCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB GTTGAATATGAGGCTCATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 343
QY 121 GGCATCATAGTGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTACATGACAGCC 180
DB GGCATCATAGTGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTACATGACAGCC 403
QY 181 ACCGTGGTTGAGAGAGCTACATCATTTGGAAGAGAGAGAGAGAGAGAGAG 240
DB ACCGTGGTTGAGAGAGCTACATCATTTGGAAGAGAGAGAGAGAGAGAGAG 463
QY 241 GGTCTAGTGGGCTCATGACACCATTTGATTTCTGAGTCTAATTTTAAAGTGGT 300
DB GGTCTAGTGGGCTCATGACACCATTTGATTTCTGAGTCTAATTTTAAAGTGGT 523
QY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGCTATGTTAGACCATTTCAAG 360
DB TTTTTCGCAAACTATGCGTTCCAAAGGATATGCTATGTTAGACCATTTCAAG 583
QY 361 ATCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB ATCTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643
QY 421 TGGGGTGCAGCAATTTCTCTGATTAAGGGGCGACATCAGCGGTATGATGAGAT 480
DB TGGGGTGCAGCAATTTCTCTGATTAAGGGGCGACATCAGCGGTATGATGAGAT 703
QY 481 GTGAACATATCGGTCATTTCTCTGCACTCATTTGCTTTTATACCTAGTGGTGG 540
DB GTGAACATATCGGTCATTTCTCTGCACTCATTTGCTTTTATACCTAGTGGTGG 763
QY 541 CTCTACTCTGTGGCATATATGATGTTGTCAGCTATTTGCACTTTTATAGAGAT 600
DB CTCTACTCTGTGGCATATATGATGTTGTCAGCTATTTGCACTTTTATAGAGAT 823
QY 601 ATCAGAGTCCCTTTGGCTGTCATCTGCAAGTACAGAGAGAGAGAGAGAGAG 660
DB ATCAGAGTCCCTTTGGCTGTCATCTGCAAGTACAGAGAGAGAGAGAGAGAG 883
QY 661 CATGCTAAATACAGAGTCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB CATGCTAAATACAGAGTCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 943
QY 721 CTGTGAATATTTCTGTATTTGATGCTGGTGAATCCCATGGAGCTTCCAGAG 780
DB CTGTGAATATTTCTGTATTTGATGCTGGTGAATCCCATGGAGCTTCCAGAG 1003
QY 781 GTCTCTCTTCAATCCCTCAGCAGCATTTGCTCAGGTCGTCCTTCTGAGAGCT 840
DB GTCTCTCTTCAATCCCTCAGCAGCATTTGCTCAGGTCGTCCTTCTGAGAGCT 1063
QY 841 TGCCTGTGATGAGCTTACCCGCAATATGATAGAGATTTGAGAGCTTCCAGAG 900
DB TGCCTGTGATGAGCTTACCCGCAATATGATAGAGATTTGAGAGCTTCCAGAG 1123
QY 901 AACCAACAGCCCTAGAGGATACAGATCCCAAGAGAGAGAGAGAGAGAGAGAT 960
DB AACCAACAGCCCTAGAGGATACAGATCCCAAGAGAGAGAGAGAGAGAGAGAT 1183
  
```

```

QY 961 CCGATCGTTCAGAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB CCGATCGTTCAGAGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
QY 1021 TTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB TTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303
QY 1081 CGGAATATACAGGTTCTCTGAGCAAAATGATCAGAGAGAGAGAGAGAGAGAG 1140
DB CGGAATATACAGGTTCTCTGAGCAAAATGATCAGAGAGAGAGAGAGAGAGAG 1363
QY 1141 ATGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB ATGAGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423
QY 1201 ACTGCTGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB ACTGCTGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
QY 1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543
QY 1321 TTTGCACTATTTCTGAGATTTACTGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB TTTGCACTATTTCTGAGATTTACTGAGAGAGAGAGAGAGAGAGAGAGAG 1603
QY 1381 TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB TTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663
QY 1441 ACTCTCTGAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB ACTCTCTGAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1723
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGTTGATGCTGCTGCTG 1560
DB CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGTTGATGCTGCTGCTG 1783
QY 1561 CACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB CACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1843
QY 1621 GAACCTGCACTGCTGAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB GAACCTGCACTGCTGAACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1903
QY 1681 GAGGCCCTCTGATGTTGATTTCCAGTCCGAGAGAGAGAGAGAGAGAGAG 1740
DB GAGGCCCTCTGATGTTGATTTCCAGTCCGAGAGAGAGAGAGAGAGAGAG 1963
QY 1741 TGA 1743
DB 1964 TGA 1966
  
```

## RESULT 9

ADV77900 standard; cDNA; 4904 BP.

ADV77900;

24-FEB-2005 (first entry)

Rat high affinity choline transporter (rCHT1) cDNA.

Choline transporter; neuromuscular disorder;  
 autonomic nervous system disorder; central nervous system disorder;  
 Parkinson's disease; Huntington's chorea; genetic disorder;  
 Alzheimer's disease; degeneration; neurological disease; schizophrenia;  
 psychiatric disease; myasthenia gravis; immune disorder; gene therapy;  
 CNS-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;

KW muscular.gen.: neuroleptic; DNA purification; gene; ss.  
 OS Rattus norvegicus.  
 FH Key Location/Qualifiers  
 FT CDS 224..1966  
 FT /\*tag= a  
 FT /product= "Rat CHT protein"  
 XX  
 XX US2004248838-A1.  
 XX  
 XX 09-DEC-2004.  
 XX  
 XX 01-DEC-2003; 2003US-00724806.  
 XX  
 XX 23-JUL-2001; 2001US-00911077.  
 XX  
 XX (UYVA-) UNIV VANDERBILT.  
 XX  
 XX Blakely RD, Apparsundaram S, Ferguson S;  
 XX  
 XX WPI; 2005-020586/02.  
 XX  
 XX P-PSDB; ADV77901.  
 XX  
 XX GENEANK; AB030947.  
 XX  
 XX  
 XX New human high affinity choline transporter (CHT) cDNA, useful for  
 PT treating a neuromuscular, autonomic or central nervous system disorder,  
 PT including Parkinson's disease, Huntington's disease, Alzheimer's,  
 PT schizophrenia.  
 XX  
 XX Example 1; SEQ ID NO 5; 73bp; English.  
 XX  
 XX The present invention provides polynucleotides encoding novel high  
 CC affinity choline transporters (CHTs), methods for their use in screening  
 CC and therapy. The invention is useful for treating neuromuscular  
 CC disorders, autonomic or central nervous system disorders such as  
 CC Parkinsons disease, Huntingtons disease, Alzheimers disease,  
 CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is  
 CC also useful in gene therapy. The present sequence is rat high affinity  
 CC choline transporter (rCHT1) cDNA. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from the US patent office at  
 CC segdata.uspto.gov/sequence.html?docID=US20040248838.  
 CC  
 CC  
 XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;  
 SQ  
 Query Match 92.7%; Score 1616.6; DB 14; Length 4904;  
 Best Local Similarity 95.5%; Pred. No. 0;  
 Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY	1	ATGCTTTTCATGTGGAAGAGCTGAGTATTATCTCTTCTACCTCTTATATTTCG	60
DB	224	ATGCTTTTCATGTGGAAGAGCTGAGTATTATCTCTTCTACCTCTTATATTTCG	283
QY	61	GTTGGAATATGGGCGCATGAGAAACCAAAACAGGGGCAACCCGAAGAGCGAGTGA	120
DB	284	GTTGGAATATGGGCGCATGAGAAACCAAAACAGGGGCAACCCGAAGAGCGAGTGA	343
QY	121	GCCATCATAGTGGGGGCGGTGACATTGGTTTGGTGGTGGTTTACATGACAGCC	180
DB	344	GCCATCATAGTGGGGGCGGTGACATTGGTTTGGTGGTGGTTTACATGACAGCC	403
QY	181	ACCTGGGTTGAGAGGCTACATCAATGGGACAGAGACAGTGTATGGGCGAGTTGT	240
DB	404	ACCTGGGTTGAGAGGCTACATCAATGGGACAGAGACAGTGTATGGGCGAGTTGT	463
QY	241	GGCTAGCTGGGCGCATGACCCCATTTGATTTCTAGCTCTATTTAGTGGTCTG	300
DB	464	GGCTAGCTGGGCGCATGACCCCATTTGATTTCTAGCTCTATTTAGTGGTCTG	523
QY	301	TTTTTTGGGAAACCTATGCGTCCAGGGATATGTACTATGTTAGACCATTCAAACAG	360
DB	524	TTTTTTGGGAAACCTATGCGTCCAGGGATATGTACTATGTTAGACCATTCAAACAG	583
QY	361	ATCTATGGAAGCCGATGAGGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGTC	420
DB	584	ATCTATGGAAGCCGATGAGGTGGGCTGCTCTTCAATCCCTGCACTGATGGAGAGATGTC	643
QY	421	TGGGCTGAGCAATTTTCTCTGCAATTTAGGGCCACCATCAGCGATCATTTGATGGAT	480
DB	644	TGGGCTGAGCAATTTTCTCTGCAATTTAGGGCCACCATCAGCGATCATTTGATGGAT	703
QY	481	GTGAACATATGCGTCATTTCTCTGCACTATGGCACTCTTTATACCTTATGGTGGG	540
DB	704	GTGAACATATGCGTCATTTCTCTGCACTATGGCACTCTTTATACCTTATGGTGGG	763
QY	541	CTTAAGCTGGGCAATATGATGATGTTTCCAGCTATTTCTGATTTTATAGACTGGG	600
DB	764	CTTAAGCTGGGCAATATGATGATGTTTCCAGCTATTTCTGATTTTATAGACTGGG	823
QY	601	ATCAGTGTCCCTTTGGCCCTGTCACATCTGACAGCACCGACATGGATTCACAGCTGG	660
DB	824	ATCAGTGTCCCTTTGGCCCTGTCACATCTGACAGCACCGACATGGATTCACAGCTGG	883
QY	661	CATGCTTAATACAGAGTCCCTGGCTGGGAACCATTTGATCAGTTGAAGCTTACCTGG	720
DB	884	CATGCTTAATACAGAGTCCCTGGCTGGGAACCATTTGATCAGTTGAAGCTTACCTGG	943
QY	721	CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAATCCATGGCAAGCTTACTCCAGAG	780
DB	944	CTTGATTAATTTTCTGTTATTTGATGCTGGGTGGAATCCATGGCAAGCTTACTCCAGAG	1003
QY	781	GTCCCTCTTCAATCCCTGACCACTATGCTCAGGACTGTGCTTCTGAGCTTTTGGG	840
DB	1004	GTCCCTCTTCAATCCCTGACCACTATGCTCAGGACTGTGCTTCTGAGCTTTTGGG	1063
QY	841	TGCCGTGATATGACTTACCCGCAATATGATAGAGGATTTGAGCTTCCACAGACTGG	900
DB	1064	TGCCGTGATATGACTTACCCGCAATATGATAGAGGATTTGAGCTTCCACAGACTGG	1123
QY	901	AACGAGACTGCTTACGGGTATCCAGATCCCAAGATTAAGAGGAAGCAGACATGATTC	960
DB	1124	AACGAGACTGCTTACGGGTATCCAGATCCCAAGATTAAGAGGAAGCAGACATGATTC	1183
QY	961	CCGATGTTTCGACATGCTGACCTGAGTATGATCATCTCTTGGGCTGGTGGTGT	1020
DB	1184	CCGATGTTTCGACATGCTGACCTGAGTATGATCATCTCTTGGGCTGGTGGTGT	1243
QY	1021	TCAGCTGTGATGCTCTCAAGCTGATCTGTCATCTGTGCGAGATTCTATGTTTGT	1080
DB	1244	TCAGCTGTGATGCTCTCAAGCTGATCTGTCATCTGTGCGAGATTCTATGTTTGT	1303
QY	1081	CGGAATATCTACAGACTTCTCTTGACAGAAATGATCAGACAAAGAAATGTGTGGTTC	1140
DB	1304	CGGAATATCTACAGACTTCTCTTGACAGAAATGATCAGACAAAGAAATGTGTGGTTC	1363
QY	1141	ATGAGGATCATCTGCTTGTGGTGGAGCATCTGCAAGCAGCATGGCTTGTGTGAGAG	1200
DB	1364	ATGAGGATCATCTGCTTGTGGTGGAGCATCTGCAAGCAGCATGGCTTGTGTGAGAG	1423
QY	1201	ACTGTGTATGGGCTGTGATCTGAGCTGTGACCTTGTCTATCATCATCATCTTCCACAG	1260
DB	1424	ACTGTGTATGGGCTGTGATCTGAGCTGTGACCTTGTCTATCATCATCATCTTCCACAG	1483
QY	1261	CTGCTCTGTGATCTTCTTCAAGAGAACCAACATCTATAGGGGAGTGTGCTTATAT	1320
DB	1484	CTGCTCTGTGATCTTCTTCAAGAGAACCAACATCTATAGGGGAGTGTGCTTATAT	1543
QY	1321	TTTGAGATTAATCTCAGAAATTAAGAGAGGACATCTTATCTTGAGAGCCCTTAATC	1380
DB	1544	TTTGAGATTAATCTCAGAAATTAAGAGAGGACATCTTATCTTGAGAGCCCTTAATC	1603
QY	1381	TTCTAACCTGTTATTAATCTTGACAAAGATGATATTAATCAATCAGAGGTTCCCATTTAA	1440
DB	1604	TTCTAACCTGTTATTAATCTTGACAAAGATGATATTAATCAATCAGAGGTTCCCATTTAA	1663

QY 1441 ACTCTCCATGGTACTCATCTTTTACCAACATTTGTTCTTATCTAGCAGATAT 1500  
 DB 1664 ACTCTCCATGGTACTCATCTTTTACCAACATTTGTTCTTATCTAGCAGATAT 1723  
 QY 1501 CTAATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATCTGTTGCGAAG 1560  
 DB 1724 CTAATTTGAAGTGAACCTTGGCTCCAAATTAATGATGATTTGATCTGTTGCGAAG 1783  
 QY 1561 CACAGTGAAGAGAACGTGACAAACCATTTAGTGAATGAATTAATTAATTAAT 1620  
 DB 1784 CACAGTGAAGAGAACGTGACAAACCATTTAGTGAATGAATTAATTAATTAATTAAT 1843  
 QY 1621 GAATTTGACACTGTGTAACCTGCGAGAGAGCTTAACCTGATTTCACTTTCAACCAATAAG 1680  
 DB 1844 GAATTTGACACTGTGTAACCTGCGAGAGAGCTTAACCTGATTTCACTTTCAACCAATAAG 1903  
 QY 1681 GAGGCTCTCTTGTGATTTGATTCAGTCCGAGAGGCTTGGGACTGAAGTAATTTACAA 1740  
 DB 1904 GAGGCTCTCTTGTGATTTGATTCAGTCCGAGAGAGCTTGGGACTGAAGTAATTTACAA 1963  
 QY 1741 TGA 1743  
 DB 1964 TGA 1966  
 RESULT 10  
 ABX94338  
 ID ABX94338 standard; cDNA; 1743 BP.  
 AC ABX94338;  
 XX  
 DT 13-JUN-2003 (first entry)  
 XX  
 DE Human cDNA encoding high affinity choline transporter, HACT.  
 XX  
 KW Human; ss; gene; HACT; high affinity choline transporter; pain;  
 KW neurotransmitter biosynthesis; learning and memory; aging; epilepsy;  
 KW neurological disorder; spasticity; myoclonus; muscle spasm;  
 KW muscle hyperactivity; stroke; head trauma; neuronal cell death;  
 KW multiple sclerosis; spinal chord injury; dystonia; Alzheimer's disease;  
 KW Myasthenia Gravis; multi-infarct dementia; AIDS dementia;  
 KW Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;  
 KW ALS; attention deficit disorder; organic brain syndrome; schizophrenia;  
 KW nicotine addiction; memory disorder; cognitive disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1743  
 FT /tag= a  
 FT /product= "HACT"  
 FT  
 XX  
 PN US6500643-B1.  
 XX  
 PD 31-DEC-2002.  
 XX  
 PF 07-SEP-2000; 2000US-00657252.  
 XX  
 PR 07-SEP-2000; 2000US-00657252.  
 XX  
 PA (UYFL ) UNIV FLORIDA.  
 XX  
 PI Wu D, Gu Y, Millard WJ, He Y;  
 XX  
 DR WPI, 2003-361355/34.  
 XX  
 DR P-PDB; AB008979.  
 XX  
 PT Novel isolated polynucleotide (1) that encodes high affinity choline  
 PT transporter protein, useful for preventing, treating or ameliorating  
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's  
 PT disease.  
 XX  
 PS Claim 2; Col 17-21; 20pp; English.

XX The invention relates to an isolated polynucleotide which encodes a high  
 CC affinity choline transporter (HACT) protein appearing as AB008979. Also  
 CC included are a polynucleotide encoding a fragment consisting of at least  
 CC about 50 amino acids of the HACT protein, a vector comprising the  
 CC polynucleotide, a composition comprising a vector comprising a  
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids  
 CC of a polynucleotide appearing as ABX94339 (encoding choline  
 CC acetyltransferase), a recombinant host cell which comprises the vector  
 CC (used to express the HACT protein or fragment). The polynucleotide is  
 CC useful as a probe or primer to detect the presence of HACT polynucleotide  
 CC in a sample, such as a biological sample, or for screening for test  
 CC agents which bind to the polynucleotide. A pharmaceutical composition  
 CC comprising the polynucleotide is useful for preventing, treating or  
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,  
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head  
 CC trauma, neuronal cell death, multiple sclerosis, spinal chord injury,  
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi-infarct  
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,  
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine  
 CC addiction, organic brain syndrome, schizophrenia or memory and cognitive  
 CC disorders. HACT is thought to be the rate limiting step in cholinergic  
 CC neurotransmitter biosynthesis and regeneration (cholinergic transmissions  
 CC are crucial to brain functions such as learning and memory). The present  
 CC sequence encodes human HACT  
 CC  
 SO Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;  
 Query Match 79.1%; Score 1378.2; DB 9; Length 1743;  
 Best Local Similarity 86.9%; Pred. No. 0;  
 Matches 1515; Conservative 0; Mismatches 228; Indels 0; Gaps 0;  
 QY 1 ATGCTCTTCCATGTGGAAGAGCTGTGATTAATCTCTTACCTCCTTAATTTCTG 60  
 DB 1 ATGCTCTTCCATGTGGAAGAGCTGTGATTAATCTCTTACCTCCTTAATTTCTG 60  
 QY 61 GTTGAATATGAGGCTGATGAAACCAAAACAGGCGAACCCAGAGAGCGAGTGA 120  
 DB 61 GTTGAATATGAGGCTGATGAAACCAAAACAGGCGAACCCAGAGAGCGAGTGA 120  
 QY 121 GCCATATATGAGGCGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180  
 DB 121 GCCATATATGAGGCGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180  
 QY 121 GCCATATATGAGGCGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180  
 DB 121 GCCATATATGAGGCGCGGTGACATTTGTTGTTGTTGTTTACATGACAGCC 180  
 QY 181 ACCTGGCTTGAAGAGAGCTATCAATATGAGAGAGAGAGCTATGAGAGAGTGT 240  
 DB 181 ACCTGGCTTGAAGAGAGCTATCAATATGAGAGAGAGAGCTATGAGAGAGTGT 240  
 QY 241 GGTCTAGCTTGGGCTCATGACACCATTTGATTAATTTAGTGTGCTG 300  
 DB 241 GGTCTAGCTTGGGCTCATGACACCATTTGATTAATTTAGTGTGCTG 300  
 QY 301 TTTTGTGAAACCTATGCGTTCCAAAGGATATGACTATGTTAAGCCATTCAACAG 360  
 DB 301 TTTTGTGAAACCTATGCGTTCCAAAGGATATGACTATGTTAAGCCATTCAACAG 360  
 QY 301 TTTTGTGAAACCTATGCGTTCCAAAGGATATGACTATGTTAAGCCATTCAACAG 360  
 DB 301 TTTTGTGAAACCTATGCGTTCCAAAGGATATGACTATGTTAAGCCATTCAACAG 360  
 QY 361 ATCTATGAAAGAGAGAGAGGAGGAGCTCTTATGCTGACCTGATGAGAGAGTTC 420  
 DB 361 ATCTATGAAAGAGAGAGGAGGAGCTCTTATGCTGACCTGATGAGAGAGTTC 420  
 QY 421 TGGGCTGAGCAATTTTCTGCTGATTAAGGAGGACCATGATGATGATGAT 480  
 DB 421 TGGGCTGAGCAATTTTCTGCTGATTAAGGAGGACCATGATGATGATGATGAT 480  
 QY 481 GTGAATATATGCTATGCTCTGACATCATTTGATTTAATCCATGAGGAGTGG 540  
 DB 481 GTGAATATATGCTATGCTCTGACATCATTTGATTTAATCCATGAGGAGTGG 540  
 QY 541 CTTACTCTGTGAGCATATGATGTTGTCAGCTATTTGCAATTTTAAAGACTGTGG 600  
 DB 541 CTTACTCTGTGAGCATATGATGTTGTCAGCTATTTGCAATTTTAAAGACTGTGG 600  
 QY 601 ATCAGTGTCCCTTTGGCTGTCATCTCTGACATCCGACATCGGATTCACAGCTGTG 660

Dp	601	ATCAGCGTCCCTTTGCAATTGTCACATCTCAGATCGCAGACATCGGGTTTCACGTCTG	660
Qy	661	CATGCTAAATACAGAGTCCCTGGCTGGGAACAATTGATCAGTTGAAGTCAACCTGG	720
Dp	661	CATGCAAAATACCAAAAGCCGTGGCTGGGAATGTGATCATCTGAAGTCTTACCTTGG	720
Qy	721	CTTGAATAATTTCTGTTATATGATGTGGGTGGAAATCCATGGCAAGCACTTTCAGAG	780
Dp	721	CTTGATATGTTTCTGTGTTGTATGTGGGTGGAAATCCATGGCAAGCACTTTCAGAG	780
Qy	781	GTCTCTCTTCATCCTCAAGCCATATGCTCAGGATCTGTCTTCTGGCAGCTTTTGGG	840
Dp	781	GTTCTCTCTTCTCTCCTCAGCCACCTATGCTCAAGTGCTGTCTTCTGGCAGCTTTGGG	840
Qy	841	TGCTGTGATAGGCTCTACCCGCCATATGATAGAGCTATTGGAGCTTCAACAGCTGG	900
Dp	841	TGCTGTGATAGGCGCATCCAGCCATCTCATTTGGGGCATTGGAGCTTCAACAGCTGG	900
Qy	901	AACCAAGCTGCTACGGGTATCCAGATCCCAAGACTAAGAGAGAGCAGACATATCTC	960
Dp	901	AACCAAGATGATATAGGGCTTCCAGATCCCAAGACTAAGAGAGAGCAGACATATTTTA	960
Qy	961	CGATCGTTCTGCAAGTACTCTGCTGGTGTATCATCTCTTTGGGCTGTGTCTGT	1020
Dp	961	CCAATGTGTTGCAAGTATCTGCGCTGTATATTTCTTCTTGGCTGTGGAGCT	1020
Qy	1021	TCAGTGTGTGATGTCTCTCAGCTGACGTGTCATCTGTGGGGAGTTCTATGTGGCT	1080
Dp	1021	TTGTGTGTGTATGTCACTACAGCAATTTCTTCATCTTGACGAAGTTCCATGTGGCA	1080
Qy	1081	CGGAATATCTACCAAGCTTTCCTTGACACAAATATGATCAGACAAGAAATGTGTGGCT	1140
Dp	1081	CGGAACATCTACCAAGCTTTCCTTGACACAAATATGATCAGACAAGAAATGTGTGGCT	1140
Qy	1141	ATGAGGATCACTGTGCTGTGTGTGGAGCACTGCAACAGCAAGCTTTGCTGACGAG	1200
Dp	1141	ATGGAATCAACAGTGTGTGTGTGGAGCACTGCAACAGCAAGCTTTGCTGACGAAA	1200
Qy	1201	ACTGTATATGAGGCTGTGGTACCTGAGCTCAGCTGACCTGTATCATCATCTTCCACAG	1260
Dp	1201	ACTGTATATGAGGCTGTGGTACCTGAGCTCAGTCTGACCTGTATCATCTTCCACAG	1260
Qy	1261	CTGCTCTGTGTACTCTTTCATCAAAAGAACCAACTTATGCGGGCAGTTGCTGTATATT	1320
Dp	1261	CTGCTCTGTGTACTCTTGTATAGGGAACCAACCTATGCGGGCAGTTGATATT	1320
Qy	1321	TTTGGACATATCCGAGAAATTAAGGAGAGAGCATATCTATCTTGACCCCTTATATC	1380
Dp	1321	TTGAGCTCTTCTCTGAGAAATTAAGGAGAGAGCATATCTGTATCTTGACCCCTTATATC	1380
Qy	1381	TTTCAACCTGTGTTATTACTGTACACAAATGTATATCAATCAAGGTTCCATTAAA	1440
Dp	1381	TTTCAACCTGTGCTATTAACCTGTAGATATATGTATATCAAGAAATTTCCATTAAA	1440
Qy	1441	ACTCTCTCAGATTAAGTCTCATTTTTCACCAATTTGTGTTCTTATCTTACGCAAGAT	1500
Dp	1441	ACACTCTCAGATGTATCAATCTTTCACCAACATTTGATCTCTATCTTACGCAAGAT	1500
Qy	1501	CTATTTTGAAGTGAACCTTGCTCCAAATTTAGATGATTTGATGTGTCGCAAG	1560
Dp	1501	CTATTTTGAAGTGAAGCTTGCTCCAAATTTAGATGATTTGATGTGTCGCAAGA	1560
Qy	1561	CACAGTGAAGGAACATGAGCAAGACCTTTAGTCAAGAAATGAATAATATCAATTAAT	1620
Dp	1561	CACAGTGAAGGAACATGATTAAGCAATTTGTCAAAATAATGAATAATTAATTAAT	1620
Qy	1621	GAACTTGAACCTGAGAACTCTGGAGAGGCTTAACCTCACTTCAACCTTACCAATTAAG	1680
Dp	1621	GAACTTGAACCTTGTGAAAGCAGACAGAGCAATGACCTCACTTCAACCTTACCAATTAAG	1680
Qy	1681	GAGGCGCTCTTGTATGTGATTCAGATCGAGGAGGCTGTGGACTGAAGATTAATTTCAA	1740

Db	1681	GAGGCGCTTCCTTGATGTTGATTCCAGTCGAGAGGGCTCGGAGCTGAAGATTATTACAA	1740
QY	1741	TGA 1743	
Db	1741	TGA 1743	
RESULT 11			
AAFB1712	ID	AAFB1712 standard; cDNA; 1743 BP.	
AAFB1712;	AC		
AAFB1712;	AC		
01-JUN-2001	DT	(first entry)	
Human high affinity choline transporter protein encoding cDNA.	DE		
High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;	XX		
ss.	KW		
Homo sapiens.	XX		
OS	XX		
Key	FT	Location/Qualifiers	
CDS	FT	1..1743	
	FT	/*tag= a	
	FT	/product= "high affinity choline transporter"	
WO200116315-A1.	PN		
08-MAR-2001.	PD		
18-AUG-2000; 2000WO-JP005545.	PP		
27-AUG-1999; 99JP-00240642.	PR		
27-DEC-1999; 99JP-00368991.	PR		
(NISC-) JAPAN SCT & TECHNOLOGY CORP.	PA		
Haga T, Okuda T;	PI		
WPI; 2001-226688/23.	DR		
P-PSDB; AAB74665.	DR		
New rat and human spinal cord high affinity choline transporters, useful	PT		
in diagnosis of Alzheimer's disease and screening promoters as drugs for	PT		
treating Alzheimer's disease.	PT		
Claim 9; Page 71-75; 90pp; Japanese.	PS		
The present sequence encodes a human (Homo sapiens) high affinity choline	CC		
transporter protein designated cho-1. The cho-1 protein has nootropic and	CC		
neuroprotective activities. The cho-1 polynucleotide and protein can be	CC		
used for the diagnosis of diseases related to the expression of cho-1 by	CC		
comparing the cho-1 polynucleotide sequence in a sample to that of a	CC		
control. Drug compositions containing the cho-1 protein or expression	CC		
promoters or inhibitors of cho-1 are useful for treating disorders	CC		
characterised by abnormal levels of cho-1, such as Alzheimer's disease	CC		
Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;	XX		
Query Match		78.9%; Score 1375; DB 4; Length 1743;	
Best Local Similarity		86.8%; Pred. No. 0;	
Matches 1513; Conservative		0; Mismatches 230; Indels 0; Gaps 0;	
1 ATGCGCTTCCATGTTGGAAGAAGCTGTGATGCTATTATTCCTCTTCACTTCTTATATTCTG	60		
1 ATGCGCTTCCATGTTGGAAGAAGCTGTGATGCTATTATTCCTCTTCACTTCTTATATTCTG	60		
61 GTTGGAAATATGCGCTGCATGTGAAACCAAAAACAGCGGCAACCCAGAAAGCGGATGAA	120		
61 GTTGGAAATATGCGCTGCATGTGAAACCAAAAACAGTGGACGCGAAGAACGCGACGCA	120		
121 GCCATTCATATGTCGGGGGCGCTGACATGTTGTTTGGTTGGTGGTTTACCATGACGCC	180		

```

Db      121 GGCATCATATGTTGGTGGCCGAGATATGGTTATTGGTTGGTGATTTACATGACAGCT 180
Qy      181 ACCCTGGGTGGAGAGGCTACATCAATGGGACAGCAAGACATGTATGGGCCAGGTTGT 240
Db      181 ACCCTGGGTGGAGAGGCTATCAATGGCACAGCTGAAGCAGTTATGTATACAGGTTAT 240
Qy      241 GGTCTAGCTTGGGGCTCATGCCCATTTGATATTTCTGAGTCTATTTTAAAGTGTCTG 300
Db      241 GGCCTAGCTTGGGGCTCAGGACCAATTTGATTTCTTAGTCTGATTTTAAAGTGTCTG 300
Qy      301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACTATGTATAGCAATTTCAAAACAG 360
Db      301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTACTATGTATAGCAATTTCAAAACAG 360
Qy      361 ATCTATGAAAAGCGCATGGGTGGGCTCTTTCATCCCTGCACTGATGGAGAGATGTTT 420
Db      361 ATCTATGAAAAGCGCATGGGTGGGCTCTTTCATCCCTGCACTGATGGAGAGATGTTT 420
Qy      421 TGGGCTGACGCAATTTTCTGCACTTAAAGGGCCACCATCAGCGTATCAATGTATGAT 480
Db      421 TGGGCTGACGCAATTTTCTGCTGTTGGAGCCACCATCAGCGTATCAATGTATGAT 480
Qy      481 GTGAACATATCGGTCAATGCTCTGCACTCAATTTGCCATTTCTTATAACCTATGGGTGG 540
Db      481 ATGCACATTTCTGTATCATCTCTGCACTCAATTTGCCATTTCTTATAACCTATGGGTGG 540
Qy      541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGGACTGTGG 600
Db      541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGGACTGTGG 600
Qy      601 ATCAGAGTCCCTTTTGGCCGTGACATCCGTCAGTACAGCAATGGGATTTCAAGCTGTG 660
Db      601 ATCAGAGTCCCTTTTGGCCGTGACATCCGTCAGTACAGCAATGGGATTTCAAGCTGTG 660
Qy      661 CATGCTAAATACCAAGTCCCTGGCTGGGAAACCATTTGAATCACTTGAAGTCTACACTGG 720
Db      661 CATGCTAAATACCAAGTCCCTGGCTGGGAAACCATTTGAATCACTTGAAGTCTACACTGG 720
Qy      721 CTTGATATTTTCTGTTATTTGATGCTGGGTGGAATCCCATGGCAAGCTTCCAGAG 780
Db      721 CTTGATATTTTCTGTTATTTGATGCTGGGTGGAATCCCATGGCAAGCTTCCAGAG 780
Qy      781 GTTCTCTCTTCTCTCTCAAGCACCCTATGCTCAAGTCTGTCTTCTTGGAGCTTTGG 840
Db      781 GTTCTCTCTTCTCTCTCAAGCACCCTATGCTCAAGTCTGTCTTCTTGGAGCTTTGG 840
Qy      841 TSCCTGTGTGATGGCTTACCCGCTATGCTATGAGACTATTTGAGCTTCCACAGACTGG 900
Db      841 TSCCTGTGTGATGGCTTACCCGCTATGCTATGAGACTATTTGAGCTTCCACAGACTGG 900
Qy      901 AACCCAGACTGCTCAAGGATATCCAGATCCCAAGACTTAAAGAGAGAGACAGCATGATTCT 960
Db      901 AACCCAGACTGCTCAAGGATATCCAGATCCCAAGACTTAAAGAGAGAGAGACAGCATGATTCT 960
Qy      961 CCGATGCTCTGAGTACCTTGGCCCTGTGTATCATCTCTTGGGGCTTGGTGTCTGT 1020
Db      961 CCGATGCTCTGAGTACCTTGGCCCTGTGTATCATCTCTTGGGGCTTGGTGTCTGT 1020
Qy      1021 TCAGTGTCTGTATGCTCTCAGCTGACTGCTCCATCTGTCCGAGCTTCTATGTTTGTCT 1080
Db      1021 TCAGTGTCTGTATGCTCTCAGCTGACTGCTCCATCTGTCCGAGCTTCTATGTTTGTCT 1080
Qy      1081 CGGAATATCTACAGCTTCTCTTCAACAAATGCAATGCAAGAAATGTTGTGGTCT 1140
Db      1081 CGGAATATCTACAGCTTCTCTTCAACAAATGCAATGCAAGAAATGTTGTGGTCT 1140
Qy      1141 ATGAGATCACTGTGCTGTTGCTCGAGCATCTGCAACAGCCATGCTTGTGCAAG 1200
Db      1141 ATGAGATCACTGTGCTGTTGCTCGAGCATCTGCAACAGCCATGCTTGTGCAAG 1200
Qy      1201 ACTGTATATGGGCTTGTGATCTGAGCTTGTCTGATCATCATCTTCCACAG 1260

```

```

Db      1201 ACTGTATATGGGCTGTGTACTGACTTCTGACCTTGTATACATGTTATCTTCCCCAG 1260
Qy      1261 CTGCTCTGTGTACTCTTCAATCAAGGAACCAACATATATGGGCAAGTTGCTGTTATTT 1320
Db      1261 CTGCTCTGTGTACTCTTGTATAGGAAACCAACATATATGGGCAAGTTGCTGTTATTT 1320
Qy      1321 TTTGACATATCTTCTGAAATTTACTGGAGAGAGCCATATCTATCTTGGAGCCCTTAATC 1380
Db      1321 TCTGGCTCTTCTGAAATTTACTGGAGAGAGCCATATCTGATCTTACGCCCTTGATC 1380
Qy      1381 TTTTACCTGTTTATTTACTCTGACAAAGATGTTATATCAATCAGAGTTCCCATTTAA 1440
Db      1381 TTTTACCTGTTTATTTACTCTGACAAAGATGTTATATCAATCAGAGTTCCCATTTAA 1440
Qy      1441 ACTCTCTGAGTATACCTCATTTCTTTACCAATTTGTTGTTCTATCTAGCAAGTAT 1500
Db      1441 ACACCTGCTGTTATCATCTATTTACCAATTTGTTGTTCTATCTAGCAAGTAT 1500
Qy      1501 CTATTTGAAAGTGAACCTTGTGCTCCAAATTTGATTTGATGCTGTGTGCAAG 1560
Db      1501 CTATTTGAAAGTGAACCTTGTGCTCCAAATTTGATTTGATGCTGTGTGCAAG 1560
Qy      1561 CACAGTGAAGAGACATGAGCAAGACATTTCTAGTCAAGAAATGAAATTTCAAATTAAT 1620
Db      1561 CACAGTGAAGAGACATGAGTAAACAATTTCTGTCAAAAATGAAATTTAAATTAAGAT 1620
Qy      1621 GAACCTTGACCTGTGAAACCTTGGGAGAGCTTAACCTGATTCATCTTCAACCAATAAG 1680
Db      1621 GAACCTTGACCTGTGAAAGCAGACAGACATGACCTTACCTCAACCTTCAACCAATAAA 1680
Qy      1681 GAGGCCCTCTGATGTTGATTTCCAGTCCGAGAGGGGTCTGGGACTGAGATAATTTTACA 1740
Db      1681 GAGGCCCTCTGATGTTGATTTCCAGTCCGAGAGGGGTCTGGGACTGAGATAATTTTACA 1740
Qy      1741 TGA 1743
Db      1741 TGA 1743

```

```

RESULT 12
AAH49207
ID AAH49207 standard; cDNA; 1743 BP.
XX
AC AAH49207;
XX
DT 26-NOV-2001 (first entry)
XX
DE Human CHOT encoding cDNA.
XX
KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;
KW neuroprotective; gene therapy; antisense therapy; degenerative disease;
KW cognitive disorder; Alzheimer's disease; ss.
XX
OS Homo sapiens.
XX
PN DE1009055-A1.
PD 30-AUG-2001.
XX
PF 28-FEB-2000; 2000DE-01009055.
XX
PR 28-FEB-2000; 2000DE-01009055.
XX
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.
XX
PI Brues M., Boenisch H;
XX
XX WPI; 2001-590709/67.
XX DR P-PDB; AAB6837.
XX
PT A new gene encoding human choline transporter, designated hCHOT is
located on chromosome 2q11-13 and is useful to treat degenerative

```



DT 15-JAN-2004 (first entry)  
XX CDNA encoding human high-affinity choline transporter (hCHT).  
XX  
XX Human: high-affinity choline transporter; hCHT; chromosome 2q12;  
XX cholinergic function; Parkinson's disease; Huntington's disease;  
XX Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;  
XX brain; cholinergic signaling; antiparkinsonian; anticonvulsant;  
XX neurotropic; neuroprotective; neuroleptic; gene; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1743  
FT /tag= a  
FT /product= "hCHT"  
XX  
XX US2003114399-A1.  
XX  
XX 19-JUN-2003.  
XX  
XX 23-JUL-2001; 2001US-00911077.  
XX  
XX 23-JUL-2001; 2001US-00911077.  
XX  
XX (BLAK/) BLAKELY R D.  
XX (APP/) APPARISUNDARAM S.  
XX (FERG/) FERGUSON S.  
XX  
XX Blakely RD, Apparsundaram S, Ferguson S;  
XX  
XX MPI; 2003-810914/76.  
XX P-PSDB; ADD50639.  
XX  
XX Novel isolated polynucleotide encoding human or mouse high affinity  
XX choline transporter polypeptide, useful in gene therapy to increase  
XX cholinergic function in a cell of a patient suffering from Alzheimer's  
XX disease.  
XX  
XX Claim 2: SEQ ID NO 1; 74pp; English.  
XX  
XX The present invention relates to the isolation of polynucleotide  
XX sequences encoding human and mouse high-affinity choline transporter  
XX (hCHT and MCHT respectively), and the proteins they encode. The gene  
XX encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
XX encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
XX polynucleotide sequence when delivered to a cell, increases cholinergic  
XX function in the cell that is in a patient having Parkinson's disease,  
XX Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
XX myasthenia gravis. The hCHT antibody is useful for controlling  
XX transporter CHT proteins to the brain, and for treating the above  
XX mentioned diseases. The antibody is also useful for diagnosing the above  
XX mentioned disorders and to detect the influence of cholinergic  
XX signaling. The present sequence encodes hCHT. Note: The sequence data  
XX for this patent was obtained in electronic format directly from the USPTO  
XX web site at seqdata.uspto.gov.  
XX  
XX Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other:  
SQ  
Query Match 78.9%; Score 1375; DB 10; Length 1743;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;  
QY 1 ATGGCTTCCATGTGAGAGAGCTGATGCTATATCTCTTCACTCCATTATTTCTG 60  
DB 1 ATGGCTTCCATGTGAGAGAGCTGATGCTATATCTCTTCACTCCATTATTTCTG 60  
QY 61 GTTGAATATGAGCTCATGAAACCAAAACAGCGGCAACCCAGAGAGCGAGTGA 120  
DB 61 GTTGAATATGAGCTCATGAAACCAAAACAGCGGCAACCCAGAGAGCGAGTGA 120  
QY 121 GCCATATATGTCGGGGCGCTGACATTTGTTTGTGGTGTGTTTACATGACGAC 180  
DB 121 GCCATATATGTCGGGGCGCTGACATTTGTTTGTGGTGTGTTTACATGACGAC 180

DB 121 GCCATATATGTCGGGGCGAGATATGTTATTTGTTGTTGATTTACATGACGCT 180  
QY 181 ACCTGGCTTGGAGAGAGCTATCATATGAGAGACGAGAGCTATGGCCAGGTTGT 240  
DB 181 ACCTGGCTTGGAGAGAGCTATCATATGAGAGAGCTATGGCCAGGTTAT 240  
QY 241 GGCTATGCTTGGGCTCATGACACCATTTGATTTCTGAGTCTAAATTTAGTGTCTG 300  
DB 241 GGCTATGCTTGGGCTCATGACACCATTTGATTTCTGAGTCTAAATTTAGTGTCTG 300  
QY 301 TTTTTCGAAACCTATGCGTTCCAAAGGATATGATCTATGTTAAGCATTTCAAACAG 360  
DB 301 TTTTTCGAAACCTATGCGTTCCAAAGGATATGATCTATGTTAAGCATTTCAAACAG 360  
QY 361 ATCTATGAAAGGAGATGGGCTGGCTCTTATGCTGACCTGATGGAGAGATGTTTC 420  
DB 361 ATCTATGAAAGGAGATGGGCTGGCTCTTATGCTGACCTGATGGAGAGATGTTTC 420  
QY 421 TGGGCTGACGAATTTCTGCTATGAGGAGCCACCATGAGCGTGTATGATGATGAT 480  
DB 421 TGGGCTGACGAATTTCTGCTATGAGGAGCCACCATGAGCGTGTATGATGATGAT 480  
QY 481 GTGAACATATCGGCTATGCTCTGCACTGATTTGCCATTTCTTTATACCTAGTGGTGG 540  
DB 481 GTGAACATATCGGCTATGCTCTGCACTGATTTGCCATTTCTTTATACCTAGTGGTGG 540  
QY 541 CTCTACTCTGAGCAATATGATGATGTTGTCAGCTATTCGATTTTATAGAGCTGTG 600  
DB 541 CTCTATCTGAGCAATATGATGATGTTGTCAGCTATTCGATTTTATAGAGCTGTG 600  
QY 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTCAGTGCAGCCAGATTCGATTCAGCTGTG 660  
DB 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTCAGTGCAGCCAGATTCGATTCAGCTGTG 660  
QY 661 CATGCTAAATPACAGAGTCTCTGCTGAGAACCATTTGAATCACTGATCTACCTGTG 720  
DB 661 CATGCTAAATPACAGAGTCTCTGCTGAGAACCATTTGAATCACTGATCTACCTGTG 720  
QY 721 CTGTATATTTCTGTTATGATGCTGGGAGAAATCCATGGCAAGCTTCCAGAGG 780  
DB 721 CTGTATATTTCTGTTATGATGCTGGGAGAAATCCATGGCAAGCTTCCAGAGG 780  
QY 781 GTCTCTCTTCAATCCACAGCACTATGCTCAAGTATGCTTCTCTGAGAGCTTTGGG 840  
DB 781 GTCTCTCTTCTTCCACAGCACTATGCTCAAGTATGCTTCTCTGAGAGCTTTGGG 840  
QY 841 TGCCTGTATGAGCTCTACCCGCAATATGATATGAGCTATTTGAGCTTCCACAGACTGG 900  
DB 841 TGCCTGTATGAGCTCTACCCGCAATATGATATGAGCTATTTGAGCTTCCACAGACTGG 900  
QY 901 AACCGAGCTGCTAGGGGATCCAGATCCCAAGATCCAAAGAGAGAGACATGATTTCTC 960  
DB 901 AACCGAGCTGCTAGGGGATCCAGATCCCAAGATCCCAAGAGAGAGACATGATTTCTC 960  
QY 961 CCGATGCTTCTGAGTACCTTGCCTGTATATCTCTTCTTTGAGTGTGCTGT 1020  
DB 961 CCGATGCTTCTGAGTACCTTGCCTGTATATCTCTTCTTTGAGTGTGCTGT 1020  
QY 1021 TGAAGTGTGATATGCTCTGAGTGAATCTGCTCATCTCTGTGGCGAGTTCTATGTTTGT 1080  
DB 1021 TGAAGTGTGATATGCTCTGAGTGAATCTGCTCATCTCTGTGGCGAGTTCTATGTTTGT 1080  
QY 1081 CGGAATATCTACAGCTTCTTCAAGAAATGATCAAGACAAAGAAATGTTGTGGTTC 1140  
DB 1081 CGGAATATCTACAGCTTCTTCAAGAAATGATCAAGACAAAGAAATGTTGTGGTTC 1140  
QY 1141 ATGAGAGTCACTGTGCTGTGTTGGAGATCTGCAACAGCATGCTTGTGACGAAG 1200  
DB 1141 ATGAGAGTCACTGTGCTGTGTTGGAGATCTGCAACAGCATGCTTGTGACGAAG 1200  
QY 1201 ACTGCTATGAGCTCTGAGTCACTGAGTCACTGATCTGATCACTATCTTCCACAG 1260  
DB 1201 ACTGCTATGAGCTCTGAGTCACTGAGTCACTGATCTGATCACTATCTTCCACAG 1260







Db	601	ATGACGGTCCCTTTGGATTGTGCATCTCTGCACTCTGCACTGCGAGACATCGGGTTACCTGCTGTG	660
Qy	661	CATGCTAAATACCAAGATCCCTGGCTGGGAACAATTGAATCAGTTGAAGTCAACCTGG	720
Db	661	CATGCGAAATACCAAAAGCCGTGGCTGGGAATGTGTGACTCATCTGAAGTCAATCTTTGG	720
Qy	721	CTGTATATTTCTTGTTATGTATGCTGGGAGGAATCCCATGGCAAGCCTACTTCAGAGG	780
Db	721	CTGTATATGTTTCTGTGTGTGTATGTCTGGGTGGAATCCATGGCAAGCACTATTCAAGG	780
Qy	781	GTCCTCTCTTCATCCTCAAGCCACTATGCTCAGGTACTGTCTTCTTGACGCTTTTGGG	840
Db	781	GTTCTCTCTTCTCTCTCAGCACTATGCTCAAGTGCTGTCTTCTTGAGCACTTTGGG	840
Qy	841	TGCTGTGGATGGCTTACCCGCCATATGCAATAGAGCTATTGGAGCTTCAACGACTGG	900
Db	841	TGCTGTGGATGGGCTATCCAGCCATCTCATTTGGGGCATTTGAGAGCAATCAACGACTGG	900
Qy	901	AACCAAGCTGCTACGGGTATCCAGATCCCAAGACTAAAGAGGAAGCAGATGATCTC	960
Db	901	AACCAAGATGATATGGGCTTCCAGATCCCAAGACTAAGAAAGGAGCAGATGATTTTA	960
Qy	961	CGATTCGTTCTGCAGTACTCTGCCCCTGTGTACATCTCTTCTTTGGGCTTGGTCTGT	1020
Db	961	CCAATTTGTTCTGCAGTATCTCTGCCCTGTGTATTTCTTCTTTGGCTTGGGCACTG	1020
Qy	1021	TCAGCTGTGTCAATGCTCCTCAGCTGCACTGCGCATCTCGTGGGGAGTGTCTATGCT	1080
Db	1021	TCGTGTGTGTATATGTCATGCAAGATTTCTTCATCTGTGAGCAAGTTCATGTGTGCA	1080
Qy	1081	CGGAATATCTACCAAGCTTTCCTTCAGACAAATGATCATGACAAGAAATGTGTGGGT	1140
Db	1081	CGGAACATCTACCAAGCTTTCCTTCAGACAAATGCTTCGACAAAGAAATGTTGGGT	1140
Qy	1141	ATGAGATCACTGTGCTGTGTGTGGAGCATCTGCAACGCCATGGCTTGTCTACGAG	1200
Db	1141	ATGGGAATCAAGTGTGTGTGTGGAGCATCTGCAAAGGCCATGGCTTGTCTACGAAA	1200
Qy	1201	ACTGTGATGAGGCTCTGGTACTGAGCTCTAGCCTGACCTGTCTACATCATCTTCCACAG	1260
Db	1201	ACTGTGATGAGGCTCTGGTACTCTAGTCTGACTTGTTCATCTGTATATCTTCCCCAG	1260
Qy	1261	CTGCTGTGTACTCTTTCATCAAGGAACAAACTTAATGGGGCAGTTGCTGTTATAT	1320
Db	1261	CTGCTGTGTGTACTCTTGTTAATGAGGAACCAACTTAATGGGGCAGTTATATGTT	1320
Qy	1321	TTTGGACATATCCGAGAAATTAATCGAGAGACATATCTATCTTGACCCCTTATC	1380
Db	1321	TTGAGCCCTTCTCTGAGAAATTAATCGAGGGAGCATATCTGATCTTCACCCCTTATC	1380
Qy	1381	TTCTAACCTGGTTTATCTCTGACAGAGATGTTATACATTCAGAGGTTCCATTTAAA	1440
Db	1381	TTCTAACCTGGCTATTACCTCGATGATATGTTATATATATCAAGAAATTTCCATTTAAA	1440
Qy	1441	ACTCTCTCATGTTACTCTTCTTCAACAATTTGTGTTCTTATCTTACCCAATAT	1500
Db	1441	ACACTGTGCAATGTTACATCAATCTTCAACCAATTTGCAATCTCTATCTAGCCAATAT	1500
Qy	1501	CTATTTTGAAATGGAAACCTTGCTCCCAAAATTAAGTGAATTTGATGCTGTGCGCAGG	1560
Db	1501	CTATTTTGAAATGGAACTTGCCACCTTAAATTAATGATGATTTGATGCTGTGCGAAGA	1560
Qy	1561	CACAGTGAGAAACATGAGCAAGCACTTCTAGTCAGAAATGAATAATCAATTAAT	1620
Db	1561	CACAGTGAGAAAACATGATTAAGCAATCTTGTCAAAAATGAATAATTAATTAAT	1620
Qy	1621	GAACTTGAACCTGAGAAACCTCGGAGAGGCTTAACCTCAGTTCAAACTTTACCAATTAAG	1680
Db	1621	GAACTTGAACCTTGTGAACCAAGCAAGCATGACCTCAGCTCAACTTTACCAATTAAG	1680
Qy	1681	GAGGCGCTCTTGTATGTTATTCAGTCCGAGGGGCTTGGGAGCTGAAGATTAATTTACAA	1740

Db	1681	GAGGCCCTTCCTGTAATGTTGATTCACATGTCAGAAAGGCTCTGGAGCTGAAGTAATTATTCAG	1743
Oy	1741	TGA	1743
Db	1741	TGA	1743
RESULT 15			
ID	ADD50646	standard; DNA, 1813 BP.	
XX	ADD50646;		
AC	ADD50646;		
XX	ADD50646;		
DT	15-JAN-2004	(first entry)	
XX			
DE	High-affinity choline transporter (CHT) associated DNA sequence #2.		
XX			
KW	High-affinity choline transporter; CHT; cholinergic function;		
KW	Parkinson's disease; Huntington's disease; Alzheimer's disease;		
KW	schizophrenia; dysautonomia; myasthenia gravis; brain;		
KW	cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;		
KW	neuroprotective; neuroleptic; ds.		
XX			
OS	Unidentified.		
XX			
FN	US2003114399-A1.		
PD	19-JUN-2003.		
XX			
XX	23-JUL-2001; 2001US-00911077.		
FP	23-JUL-2001; 2001US-00911077.		
PR	23-JUL-2001; 2001US-00911077.		
XX			
XX	(BLAK/) BLAKELY R D.		
PA	(APPA/) APPARUNDARAM S.		
PA	(FERG/) FERGUSON S.		
XX			
XX	Blakely RD, Apparsundaram S, Ferguson S;		
XX	WPI; 2003-810914/76.		
DR			
XX			
PT	Novel isolated polynucleotide encoding human or mouse high affinity		
PT	choline transporter polypeptide, useful in gene therapy to increase		
PT	cholinergic function in a cell of a patient suffering from Alzheimer's		
PT	disease.		
XX			
PS	Disclosure; SEQ ID NO 9; 74pp: English.		
XX			
CC	The present invention relates to the isolation of polynucleotide		
CC	sequences encoding human and mouse high-affinity choline transporter		
CC	(hCHT and mCHT respectively), and the proteins they encode. The gene		
CC	encoding hCHT is located on chromosome 2q12. The polynucleotide sequence		
CC	encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT		
CC	polynucleotide sequence when delivered to a cell, increases cholinergic		
CC	function in the cell that is in a patient having Parkinson's disease,		
CC	Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or		
CC	myasthenia gravis. The hCHT antibody is useful for controlling		
CC	transporter CHT proteins to the brain, and for treating the above		
CC	mentioned diseases. The antibody is also useful for diagnosing the above		
CC	mentioned disorders and to detect the influence of cholinergic		
CC	signalling. The present DNA sequence of unknown function is provided in		
CC	the electronic sequence data but is not mentioned in the printed		
CC	specification. Note: the sequence data for this patent was obtained in		
CC	electronic format directly from the USPTO web site at seqdata.uspto.gov.		
XX			
XX	Sequence 1813 BP; 440 A; 406 C; 417 G; 550 T; 0 U; 0 Other;		
XX			
Query Match	78.9%; Score 1375; DB 10; Length 1813;		
Best Local Similarity	86.8%; Pred. No. 0;		
Matches 1513; Conservative	0; Mismatches 230; Indels 0; Gaps 0;		
Oy	1 ATGCGTTTCATGTGAGAGACTGGTATGCTATTAATCTCTTCACTCTTATTAATTTCTG	60	

Db 19 ATGCTTCCATGAGAGACTGATAGCTATCATGCTGCTACCTTCTAATTTTCTG 78  
QY 61 GTTGAATATGAGCTGATGAAAAACAAAAACAGCGCAACCCAGAAAGCGCAGTGA 120  
Db 79 GTTGAATATGAGCTGCTCGAGAAACAAAAACATGGCAGCGAGAAAGCGCAGCA 138  
QY 121 GCCATCATAGTGGGGCCGTGACATTTGTTTGGTTGGTTTAAACATGACAGCC 180  
Db 139 GCCATCATAGTGGGGCCGATATTTGTTTATTTGGTTGGATTTTACATGACAGCT 198  
QY 181 ACCTGGTTGAGAGAGCTACATCATGGGACAGAGAAAGAGTGAATGGCGAGTTGT 240  
Db 199 ACCTGGTTGAGAGAGGCTATCATATGACACAGCTGAAGCATTATATACAGTTAT 258  
QY 241 GGTCTAGCTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTCTG 300  
Db 259 GGCCTAGCTGGGCTCAGGCAACCAATTTGATATTTCTTAACTGATTTTAAAGTGGCTG 318  
QY 301 TTTTTCGGAACCTATGCTTCCAAAGGATATGATCTATGTTTGAACCATTTCAAACG 360  
Db 319 TTTTTCGGAACCTATGCTTCCAAAGGATATGATCTATGTTTGAACCATTTCAAACG 378  
QY 361 ATCTATGAAAAAGCGATGGGAGGCTCTTATCCCTGCACTGATGGGAGAGATTTTC 420  
Db 379 ATCTATGAAAAAGCGATGGGAGGCTCTTATTTATTTCTGCACTGATGGGAGAAATGTTTC 438  
QY 421 TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACCATCAGCGTATCATTTGATGTGAT 480  
Db 439 TGGGCTGAGCAATTTTCTGCTTTGGGAGCCACATCAGCGTATCATTTGATGTGAT 498  
QY 481 GTGAACATATCGGTATGTTCTCTGCACTATTTGCTTTTAAACCTATGGGTGGG 540  
Db 499 ATGCAATTTCTGTATCATCTCTGCACTATTTGCACTCTGTACACATCGTGTGGAGGG 558  
QY 541 CTCTACTCTGAGCAATATATGATGTTGTCAGCTATTTCTGATTTTAAAGACTGGG 600  
Db 559 CTCTATTTCTGTGGCTTACATGATGTTGCTGCTCTTTGCACTTTTGTAGGGCTGTGG 618  
QY 601 ATCAGTGTCCCTTTTGGCCCTGTACATCTCTGCACTCAGCAATCGGATTTCAAGCTGTG 660  
Db 619 ATCAGTGTCCCTTTTGGCAATGTGCAATCTCTGCACTCAGCAATCGGATTTCAAGCTGTG 678  
QY 661 CATGCTAATATCAAGAGTCCCTGGCTGGGAAACATTTGAATCAAGTTGATCTACACTGG 720  
Db 679 CATGCTAATATCAAGAGTCCCTGGCTGGGAACTGTTGACTCATCTGAAGTCTTACTTTGG 738  
QY 721 CTTGATTAATTTCTGTTATTTGATGCTGGGTGAATCCCATGCAAGCCACTTCCAGAGG 780  
Db 739 CTTGATTAATTTCTGTTATTTGATGCTGGGTGAATCCCATGCAAGCACTTTCCAGAGG 798  
QY 781 GTCTCTCTTCATCTCTCAGCAACCTATAGCTAGGTATGCTTCTCTGGCAGCTTTTGGG 840  
Db 799 GTTCTCTCTTCTCTCAGCAACCTATAGCTAGGTATGCTTCTCTGGCAGCTTTTGGG 858  
QY 841 TGGCTGGTGAATGGCTCTACCCGCCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900  
Db 859 TGGCTGGTGAATGGCTCTACCCGCCATATCTCATTTGGGCACTTGGAGCATCAACAGACTGG 918  
QY 901 AACCAAGATGCTACGGGATCCAGATCCCAAGACTAAGAGAGAAAGCAGACATGATCTC 960  
Db 919 AACCAAGATGCTACGGGCTTCCAGATCCCAAGACTAAGAGAGAGCAGACATGATTTTA 978  
QY 961 CCGATGCTTTCTGCACTCTGCTGCTGTGATCATCTCTTCTTTGGGCTTGTGCTGT 1020  
Db 979 CCAATGTTCTGCACTATCTGCTGCTGTATATTTCTTTCTTTGGTCTTGTGAGATT 1038  
QY 1021 TCAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Db 1039 TCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098  
QY 1081 CGGAATATCTACCACTTTCTCTTCAAGCAAAATGATCAGACAGGAATTTGTGGGCTC 1140  
Db 1099 CGGAATATCTACCACTTTCTCTTCAAGCAAAATGCTTGGGACAAAGAAATCGTTTGGGTT 1158

QY 1141 ATGAGATCACTGCTGCTTGTGTGTGAGACATCTGCAACAGCCATGCTTTGTGAGAG 1200  
Db 1159 ATGCAATCAAGTGTGTGTGTGTGAGACATCTGCAACAGCCATGCTTTGTGAGAGAA 1218  
QY 1201 ACTGTGATGGGCTTGT 1260  
Db 1219 ACTGTGATGGGCTTGT 1278  
QY 1261 CTGCTGTGTATCTCTTCAATGAAGAACCAACACTTATGGGGCAGTTGCTGTATATT 1320  
Db 1279 CTGCTGTGTATCTCTTGT 1338  
QY 1321 TTTGAGCTATTTCTGAGATTAATGAGAGAGCCATATCTATCTGACGCTTTAATC 1380  
Db 1339 TCTGGCTCTTCTGAGATTAATGAGAGAGCCATATCTGTATCTTCAAGCCTTGATC 1398  
QY 1381 TTTCTACCTGTGTATTAATCTGACAAAGATGTATATCATCAAGGTTCCCATTTTAA 1440  
Db 1399 TTTCTACCTGTGTATTAATCTGATGATATATATATCAAGAAATTTCCATTTTAA 1458  
QY 1441 ACTCTCTCAATGCTTACCTCAATTTCTTTACCAACATTTGTTTCTTATCTAGCCAGTAT 1500  
Db 1459 ACATCTGCAATGCTTACATCAATTTCTTTAACCAACATTTGATCTCTATCTAGCCAGTAT 1518  
QY 1501 CTATTTGAAAGTGAACCTTGGCTTCAAAATTTAGATGATTTGATGCTGTGTGCAAGG 1560  
Db 1519 CTATTTGAAAGTGAACCTTGGCTTCAAAATTTAGATGATTTGATGCTGTGTGCAAGG 1578  
QY 1561 CACAGTGAAGAAATGAGAGCAAGCCATTTCTAGTCAAGAAATGAAATTAATTAAT 1620  
Db 1579 CACAGTGAAGAAATGAGAGCAATTTCTTGTCAAAATGAAATTAATTAATTAAT 1638  
QY 1621 GAACTTGCACTGTGAAACCTCGGAGAGCTTAACTCAGTTCAACTTTCAACAAATAAG 1680  
Db 1639 GAACTTGCACTGTGAAACCTCGGAGAGCAAGCACTGAGCTCAGCTCAACTTTCAACAAATAAG 1698  
QY 1681 GAGGCTCTCTTGTATGTTGATTTCCAGTCCGAGAGGGGTCTGGACTGAAGATTAATTAAG 1740  
Db 1699 GAGGCTCTCTTGTATGTTGATTTCCAGTCCGAGAGGGGTCTGGAGCTGAAGATTAATTAAG 1758  
QY 1741 TGA 1743  
Db 1759 TGA 1761

Search completed: July 11, 2006, 00:09:53  
Job time : 1053 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: July 10, 2006, 14:02:30 ; Search time 855 Seconds  
(without alignments)  
7094.569 Million cell updates/sec

Title: US-10-724-806-4

Sequence: 1 MFPHVEGVAILFYLLIFL.....FALLDVDSPEGSGETEDNLQ 580

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues  
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-O=/abs/ABBSWEB.spool/US10724806/runat.10072006.140239.7586/app\_query.fasta.1  
-DB=N.Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs050p  
-USER=US10724806 @CCN 1.1.1147 @runat.10072006.140239.7586 -NCRP=6 -ICPU=3  
-NO MMAP -NRC SCORES=0 -WAIT -DSYBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.8:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2004as:\*  
12: geneseqn2004bs:\*  
13: geneseqn2004cs:\*  
14: geneseqn2005s:\*  
15: geneseqn2006s:\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2993	100.0	1743	10	ADD50640 CDNA enco
2	2993	100.0	1743	10	ADD50660 CDNA enco
3	2993	100.0	1743	14	ADV77898 Mouse hig

4	2993	100.0	1743	14	ADV77918	Adv77918 Mouse hig
5	2968	99.2	1743	4	AAFB1713	AaFB1713 Mouse hig
6	2967	99.1	4938	5	AD002457	AaD002457 Mouse p4p
7	2956	98.8	1743	4	AAFB1711	AaFB1711 Rat hig
8	2956	98.8	4304	10	ADD50642	ADD50642 CDNA enco
9	2956	98.8	4304	14	ADV77900	Adv77900 Rat hig
10	2795	93.4	1743	4	AAFB1712	AaFB1712 Human hig
11	2795	93.4	1743	5	AAH49207	AaH49207 Human CHO
12	2795	93.4	1743	9	ABX94338	AbX94338 Human CDN
13	2795	93.4	1743	10	ADD50638	ADD50638 CDNA enco
14	2795	93.4	1743	14	ADV77896	Adv77896 Human hig
15	2795	93.4	1813	10	ADD50646	ADD50646 High-affi
16	2795	93.4	1813	14	ADV77904	Adv77904 Human hig
17	1491.5	49.8	1833	12	ADV93831	Adv93831 CDNA enco
18	1444.5	48.3	1731	4	AAFB1710	AaFB1710 C. elegans
19	1444.5	48.3	1985	10	ADD50644	ADD50644 High-affi
20	1444.5	48.3	1985	14	ADV77902	Adv77902 Nematode
21	1413	47.2	4223	4	ABL29568	AbL29568 Drosophila
22	1405	46.9	1729	6	ABL29569	AbL29569 Drosophila
23	1099.5	36.7	1461	6	ABK73210	AbK73210 Bacillus
24	1096	36.6	8760	5	AAH49206	AaH49206 Human CHO
25	1096	36.6	26323	14	AEB16466	AeB16466 Human sol
26	1096	36.6	119040	10	ADD50656	ADD50656 BAC seque
27	1096	36.6	119040	14	ADV77914	Adv77914 Human BAC
28	1096	36.6	142299	10	ADD50651	ADD50651 BAC seque
29	1096	36.6	142299	14	ADV77909	Adv77909 Human BAC
30	601	20.1	386	5	AA002461	Aa002461 Mouse p4p
31	544	18.2	329	12	ADF08298	AdF08298 Subacute
32	342	11.4	349980	5	AAFB6431	AaFB6431 Pyrococcus
33	318	10.6	2475	6	AB212050	Ab212050 Human pol
34	318	10.6	2475	12	ADMA4568	AdMA4568 Novel hum
35	313	10.5	2028	8	ABL57735	AbL57735 Human sbg
36	313	10.5	2105	8	ABT33401	AbT33401 NOYX DNA
37	313	10.5	2105	12	ADH42364	AdH42364 Novel hum
38	313	10.5	2105	12	ADH42358	AdH42358 Novel hum
39	313	10.5	2200	4	AAE27713	AaE27713 Human tura
40	313	10.5	2384	13	AD089141	Ad089141 Human tura
41	313	10.5	2384	13	ADR44927	AdR44927 Polypepti
42	312	10.4	2028	4	AAFB4022	AaFB4022 Novel hum
43	312	10.4	2456	4	AAFB4023	AaFB4023 Novel hum
44	310	10.4	2326	6	ABK10764	AbK10764 Human sod
45	307	10.3	2043	13	AD089049	Ad089049 Human uro

## ALIGNMENTS

RESULT 1  
ADD50640  
ID ADD50640 standard; CDNA; 1743 BP.  
XX  
AC ADD50640;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE CDNA encoding mouse high-affinity choline transporter (mCHT) #1.  
XX  
KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;  
XX Parkinson's disease; Huntington's disease; Alzheimer's disease;  
XX schizophrenia; dysautonomia; myasthenia gravis; brain;  
KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
KW neuroprotective; neuroleptic; gene; ss.  
XX  
OS Mus sp.  
XX  
FH  
FT Key Location/Qualifiers  
FT CDS 1..1743 /tag= a  
FT /product= "mCHT #1"  
PN US2003114399-A1.  
XX  
PD 19-JUN-2003.  
XX

PF 23-JUL-2001; 2001US-00911077.  
XX 23-JUL-2001; 2001US-00911077.  
XX 23-JUL-2001; 2001US-00911077.  
XX (BLAK/) BLAKELY R D.  
PA (APPA/) APPASUNDARAM S.  
PA (FERG/) FERGUSON S.  
XX  
XX Blakely RD, Appasundaram S, Ferguson S;  
XX MPI: 2003-810914/76.  
DR P-PSDB; ADD50641.  
XX  
XX Novel isolated polynucleotide encoding human or mouse high affinity  
PT choline transporter polypeptide, useful in gene therapy to increase  
PT cholinergic function in a cell of a patient suffering from Alzheimer's  
PT disease.  
XX  
XX Claim 30; SEQ ID NO 3; 74pp; English.  
XX  
XX The present invention relates to the isolation of polynucleotide  
CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic  
CC function in the cell that is in a patient having Parkinson's disease,  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter/CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signaling. The present sequence encodes mCHT. Note: The sequence data  
CC for this patent was obtained in electronic format directly from the USPTO  
CC web site at seqdata.uspto.gov.  
XX  
XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	6.05e-314	Length:	1743
Score:	2993.00	Matches:	580
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-724-806-4 (1-580) X ADD50640 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
DB 1 ATGCCTTTCCATGTGGAGAGACTGCTAGCATATTATCTCTTCACTCTTAAATTTCTG 60  
QY 21 ValGlyIleTrrAlaIleTrrLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40  
DB 61 GTTGGAAATGGGCTGCAATGAAACCAAAACAGCGGCAACCCAGAAAGCGCACTGA 120  
QY 41 AlaIleIleValaGlyGlyArgAspIleGlyLeuLeuValaGlyGlyPheThrMetThrAla 60  
DB 121 GCCATCATATGTGGGGGCGGTGACATTGGTTGTTGGTGGTGTATACATGACAGCC 180  
QY 61 ThrTrrValaGlyGlyGlyTyrIleasnGlyThrAlaGluAlaValaTyrGlyProGlyCys 80  
DB 181 ACCTGGGTTGGAGAGAGCTACATCAATGGAGACAGAGCACTGTATGGGCGAGTTGT 240  
QY 81 GlyLeuAlaTrrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
DB 241 GGCTTAGCTTGGGCTATGACACCATTTGATTTCTGTAGTCTAATTTTAGGTGGTCTG 300  
QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120  
DB 301 TTTTTCGCAAAACCTATGGCTTCCAGAGGATATGTGACATATGTAGACCATTCAAACG 360  
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140

DB 361 ATCTATGGAAAGCGATGGGTGGGCTGCTTTCATCTCGCATGATGGAGAGTGTTC 420  
QY 141 TrpAlaAlaIlePheSerAlaLeuGlyValaThrIleSerValIleIleAspValaAsp 160  
DB 421 TGGGCTGACGAATTTTCTTCGATTTAGGGGCCACCATCAGCGTATCATTTGATGTGAT 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
DB 481 GTGAACATATGGCTCATGTCTTCGACACTTGGCATTTGATTAACCTTATGGTGGTGG 540  
QY 181 LeuTyrSerValAlaTyrThrAspValaGlnLeuPheCysIlePheIleGlyLeuTrrp 200  
DB 541 CTCTACTGTGGCATATATGATGATGTTCAGGCTATTTCTGATTTTATAGACTGTGG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
DB 601 ATCATGTGCTCTTTGGCCCTGACATCTCGACATCTCGACATCGCATCGACTGTG 660  
QY 221 HisAlaLysTyrGlnSerProTrrpLeuGlyThrIleGlySerValGluValTyrThrTrp 240  
DB 661 CATGCTTAATACAGAGTCCCTGGGGAACCATTTGAATCAGTTGAAGTCTACACCTGG 720  
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTrrpGlnAlaTyrPheGlnArg 260  
DB 721 CTGTAAATTTTCTGTATATGATGCTGGGTGGAATCCATGCAAGCTTACTTCCAGAG 780  
QY 261 ValLeuSerSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
DB 781 GTCCCTCTTCATCTTCAGCACCATATGCTCAGGATAGTCTCTGCTGGACGCTTTGGG 840  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrSptr 300  
DB 841 TGCCGTGGATGGCTCTACCCGCCATATGAGATGAGATTTGGAGCTTCCACAACATCG 900  
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluAlaAspMetIleLeu 320  
DB 901 AACCAAGCTGCTAGGGATCCAGATCCCAAGACTTAAGAGAAAGCAAGCATGATTTCTC 960  
QY 321 ProIleValLeuGlnTyrLeuGlyProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
DB 961 CCGATCGTTTCGACATTACTCTGCTGCTGTATCATCTCTTGGGCTGTGCTGT 1020  
QY 341 SerAlaAlaValaMetSerSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
DB 1021 TCAGCTGTGTCATGTCTCAGCTGACTGTCATCTGTCATCTGTCGAGATTTGTTGCT 1080  
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrrpAl 380  
DB 1081 CGGAATATCTACAGCTTTCTTCAGACAAATGATCAGACAGAAATTTGTGGGTTC 1140  
QY 381 MetArgIleThrValLeuValaPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
DB 1141 ATGAGGATCAGCTGTGCTGTGTGGAGATCTGCAGACGACATGCGCTTGTGTCAGAG 1200  
QY 401 ThrValTyrGlyLeuTrrpTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
DB 1201 ACTGTGTATGGGCTCTGGAGCTGAGCTCTGCTGTCTATCATATCACTTCCACAG 1260  
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValaIleGlyTrrle 440  
DB 1261 CTGCTGTGTACTCTTCTCAAGAGAACCAACACTTATGGGCGATTTGCTGTTATAT 1320  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyProTyrLeuTyrLeuGlnProLeuIle 460  
DB 1321 TTTGACATATTTCTGAGATTAATCTGAGAGAGCCATATCTATCTGACGCCCTTAATC 1380  
QY 461 PheTyrProGlyTyrTrrpSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
DB 1381 TTCTACCCCTGGTTATTAATCTGACAAAGATGTATATCAACAAAGGTTCCCATTTAA 1440  
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500

Db 1441 ACTCTCTCATGGTGTACTCTCATCTTTACCAACATTTGTGTTTCTTATCTAGCCAAAGTAT 1500  
QY 501 LeuPheGluSerGlyThrLeuProProlySLeuAspValPheAspAlaValAlaIaArg 520  
Cc 1501 CTAATTTGAAAGTGAACCTTGCTCCAAATTTGATGATTTTGAATCTGTTGTCGAAAG 1560  
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
Cc 1561 CACAGTGAAGAGAAACATGACCAAGACCATTTCTAGTCAGAAATGAAATATCAAAATTTAAAT 1620  
QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
Cc 1621 GAACCTTGCACTGTGTAACCTCGGACAGACCTTAACCTCACTTCAACCAATTAAG 1680  
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
Cc 1681 GAGGCCCTCTGATGTGATTTCCAGTCCGAGGGGCTGCGACCTGAAGATTAATTTACAA 1740  
RESULT 2  
ADD50660  
ID ADD50660 standard; cDNA; 1743 BP.  
XX  
AC ADD50660;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE cDNA encoding mouse high-affinity choline transporter (mCHT) #2.  
XX  
KW Mouse; high-affinity choline transporter; mCHT; cholinergic function;  
KW Parkinson's disease; Huntington's disease; Alzheimer's disease;  
KW schizophrenia; dysautonomia; myasthenia gravis; brain;  
KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
KW neuroprotective; neuroleptic; gene; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT 1.1743  
FT CDS /tag= a  
FT /product= "mCHT #2"  
XX  
XX US2003114399-A1.  
XX  
PN 19-JUN-2003.  
XX  
PD 23-JUL-2001; 2001US-00911077.  
XX  
PR 23-JUL-2001; 2001US-00911077.  
XX  
XX (BLAK/) BLAKELY R D.  
XX PA (APPA/) APPARISUNDARAM S.  
XX PA (FERG/) FERGUSON S.  
XX  
PI Blakely RD, Apparsundaram S, Ferguson S;  
PI  
DR MPI; 2003-810914/76.  
XX P-PSDB; ADD50661.  
XX  
XX Novel isolated polynucleotide encoding human or mouse high affinity  
XX choline transporter polypeptide, useful in gene therapy to increase  
XX cholinergic function in a cell of a patient suffering from Alzheimer's  
XX disease.  
XX  
XX Example 4; SEQ ID NO 23; 74bp; English.  
XX  
XX The present invention relates to the isolation of polynucleotide  
XX sequences encoding human and mouse high-affinity choline transporter  
XX (hCHT and mCHT respectively), and the proteins they encode. The gene  
XX encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
XX encoding hCHT, is useful for expressing hCHT recombinantly. The hCHT  
XX polynucleotide sequence when delivered to a cell, increases cholinergic  
XX function in the cell that is in a patient having Parkinson's disease,  
XX Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or

Cc myasthenia gravis. The hCHT antibody is useful for controlling  
Cc transporter CHT proteins to the brain, and for treating the above  
Cc mentioned diseases. The antibody is also useful for diagnosing the above  
Cc mentioned disorders and to detect the influence of cholinergic  
Cc signalling. The present sequence encodes mCHT. Note: The sequence data  
Cc for this patent was obtained in electronic format directly from the USPTO  
Cc web site at [seqdata.uspto.gov](http://seqdata.uspto.gov).  
XX  
SQ Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 6,05e-314 Length: 1743  
Score: 2993.00 Matches: 580  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 10 Gaps: 0  
US-10-724-806-4 (1-580) x ADD50660 (1-1743)  
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyreLeuLeuIlePheLeu 20  
Db 1 ATGCTTTTCATGTGGAAGACTGTGATGATATCTCTTCTACCTCTTATATTTCTG 60  
QY 21 ValGlyIleTyrAlaAlaTyrLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40  
Db 61 GTTGGAAATATGGGCTGCATGAAACCAAAACAGGGGCAACCCAGAGAGCGCAGTGA 120  
QY 41 AlaIleIleValAlaGlyValArgAspIleGlyLeuLeuValAlaGlyPheThrMetTyrAla 60  
Db 121 GCCATATATGTGGGGGCGGTGACATGTGTTGTTGGTGGTGGTTTACCAAGACGCC 180  
QY 61 ThrTyrValGlyValGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
Db 181 ACTGGGTTGAGAGAGGCTACATCAATGAGACAGCAAGACAGTATGGGCCAGGTTGT 240  
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeu 100  
Db 241 GGTCTAGCTTTGGGCTCATGACCCATTTGATATCTGAGTCTAATTTTAAAGTGCTG 300  
QY 101 PhePheAlaLysProMetLysSerLysGlyTyrValThrMetLeuAspProPheLysGln 120  
Db 301 TTTTTCGCAAAACCTATGCGTTCGCAAGGATATGACTATGTAGACCATTCAAACAG 360  
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140  
Db 361 ATCTATGAAAGCGCATGGGCTGGCTCTTCTATCTCTGACTGATGGAGAGATGTTTC 420  
QY 141 TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
Db 421 TGGGCTGCAACATTTTCTCTGCAATTAAGGGCCACCATCAGCGTATCATATGATG 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
Db 481 GTGAAACATATCGGTCAATGCTCTGCAATCATTCATTCCTTATATCCCTAGGGGTGG 540  
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
Db 541 CTCTACTCTGTGCATATCTGATGTGTCACACTATTCGCAATTTTATAGACTGTGG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
Db 601 ATCAGTGCCTCTTTGCCCTGTCACATCTGCAGTCACCGACATCGGATTCACAGCTGTG 660  
QY 221 HisAlaLysTyrGlnSerProTyrPheLysThrIleGluSerValGluValTyrThrTyr 240  
Db 661 CAGCTTAATATACAGAGTCCCTGGCGGAGACCATTTGAATGACTTACACCTGG 720  
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260  
Db 721 CTTGATATTTTCTGATATGATGCTGGGTGGAATCCATGGCAACCTTCCAGAGG 780  
QY 261 ValLeuSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280



QY ThrTrpValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
 Db ACCTGGGTTGGAGAGGCTACATCAATGGGACAGCAAGCACTGATGGGCGAGTTGT 240  
 QY GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
 Db GGCTGAGCTTGGGCTCATGACCACTGGATATTCTGTGAGCTTAATTTAGGTGGTCG 300  
 QY PhePheAlaLysProMetLysSerLysGlyTyrValThrMetLeuAspPropheLysGln 120  
 Db TTTTTCGCAAAACCTATGGCTTCCAAAGGATATGATGATGATTAAGTAAACCATTCAAACG 360  
 QY IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
 Db ATCTATGAAAGGCGATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTTC 420  
 QY TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db TGGGCTGCAGCAATTTTCTCTGCAATTAGGGCCACCATCAGCGTGAATCATTTGATGGAT 480  
 QY ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
 Db GTGAACATATCGGTCAATGCTCTGCACTCATTCATTCCTTAATCCCTAGTGGGTGGG 540  
 QY LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
 Db CTCTACTCTGTCGCAATACATGATGATGTCACACTTCTGCACTTTTATAGACCTGTGG 600  
 QY IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 Db ATCAGAGTCCCTTTGCTCCCTGTCACATCCGCACTGACGACGACATCCGATTCACAGCTGG 660  
 QY HisAlaLysTyrGlnSerProTyrPleuGlyThrIleGlySerValGluValTyrThrTyr 240  
 Db CATGCTAAATATCCAGAGTCCCTGGCTGGGAACCATTTGAATCAATTAAGCTACACCTGG 720  
 QY LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260  
 Db CTTGATATATTTTCTGTTATGATGCTGGTGGATCCCATGGAGGCTTATCCAGAGG 780  
 QY ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
 Db GTCCTCTCTTCATCCCTCAGGCACTTAATGCTCAGTACTGCTTCTCCGAGCTTTTGGG 840  
 QY CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
 Db TGCCTGGTATGGCTCTACCCCGCATATGATGAGCTATTGGAGCTTCCACAGACTGG 900  
 QY AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320  
 Db AACCAACACTGGCTACGGGATCCAGATCCCAAGACTTAAGAGAGAGAGACATGATTCCTC 960  
 QY ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
 Db CGGATGCTTCTGAGTACCTCTGCTGCTGCTATCATCTCTTCTTGGGCTTGGTGGCTGT 1020  
 QY SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
 Db TCAGCTGCTGCTCTGCTGCTCAGCTGCTCATCTGCTGCTGCTGCTGCTTATGTTTGGCT 1080  
 QY ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrPval 380  
 Db CGGAATATCTACACGCTTCTCTTCCAGCAAAATGCACTCAGACAAAGAAATGTGTGGTTC 1140  
 QY MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
 Db ATAGAGATCACTGCTGTGTGTGTGGAGCATCTGCAACACCATGCTTGTGTCAGCAAG 1200  
 QY ThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
 Db ACTGTATATGGGCTGTGGTACCTGAGCTGTGACCTTGTCTAATCATCATCTTCCACAG 1260

QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
 Db 1261 CTGCTCTGTGTAATCTCTTCAAAAGAACCAACACTTAATGGGCACTGTGCTTATATT 1320  
 QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460  
 Db 1321 TTTTGACTATCTCTCGAAATTAATCTGGAGAGAGCCATATCTATCTTATGCAAGCCCTTAATC 1380  
 QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPhePropheLys 480  
 Db 1381 TTTTACCTGGTATTAATCTGACAAAGATGATATACATCAATCAGAGGTTCCCATTTAA 1440  
 QY ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
 Db 1441 ACTCTCTCAATGGTTATACCTCATCTTTTACCAACATTTGTGTTCTTATCTAGCCAAATAT 1500  
 QY 501 LeuPheGlySerGlyThrLeuProPolysLeuAspValPheAspAlaValAlaArg 520  
 Db 1501 CTATTTGAAAGTGGAACCTTGCTCCAAATTAATGATGATTTGATGATCTGTGTGCGAAG 1560  
 QY HisSerGluGlnAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn 540  
 Db 1561 CACAGTGAAGAGAACATGGACAAAGACCATTTCTAGTCAAAATGAAATATCAATTAAT 1620  
 QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
 Db 1621 GAACCTGCACCTGTGAACCTGGGAGAGCTTACCTCAGTTCAACTTTCAACCAATTAAG 1680  
 QY 561 GluAlaLeuLeuAspValAspSerSerProGlyLysGlySerGlyThrGluAspAsnLeuGln 580  
 Db 1681 GAGGCCCTCTGATGATGATTCACAGTCCGAGGGGCTGTGGAGACTGAAGATTAATTACAA 1740

RESULT 4  
 ADV77918  
 ID ADV77918 standard; cDNA, 1743 BP.  
 XX  
 AC ADV77918;  
 XX  
 DT 24-FEB-2005 (first entry)  
 XX  
 DE Mouse high affinity choline transporter (mCHT) cDNA.  
 XX  
 KW Choline transporter; neuromuscular disorder;  
 KW autonomic nervous system disorder; central nervous system disorder;  
 KW parkinsons disease; huntingtons chorea; genetic disorder;  
 KW alzheimers disease; degeneration; neurological disease; schizophrenia;  
 KW psychiatric disease; myasthenia gravis; immune disorder; gene therapy;  
 KW cms-gen.; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;  
 KW muscular-gen.; neuroleptic; DNA purification; gene; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1743  
 FT /\*tag= a  
 FT /product= "Mouse CHT protein"  
 XX  
 PN US2004248838-A1.  
 XX  
 PD 09-DEC-2004.  
 XX  
 PF 01-DEC-2003; 2003US-00724806.  
 XX  
 PR 23-JUL-2001; 2001US-00911077.  
 XX  
 PA (UYVA-) UNITV VANDERBILT.  
 XX  
 PI Blakely RD, Apparsundaram S, Ferguson S;  
 XX  
 DR WPI; 2005-020586/02.  
 DR P-PsDB; ADV77919.  
 XX  
 PT New human high affinity choline transporter (CHT) cDNA, useful for

PT treating a neuromuscular, autonomic or central nervous system disorder,  
PT including Parkinson's disease, Huntington's disease, Alzheimer's,  
PT schizophrenia.

PS Example 4: SEQ ID NO 23; 73pp; English.

XX The present invention provides polynucleotides encoding novel high  
CC affinity choline transporters (CHTs), methods for their use in screening  
CC and therapy. The invention is useful for treating neuromuscular  
CC disorders, autonomic or central nervous system disorders such as  
CC Parkinsons disease, Huntingtons disease, Alzheimers disease,  
CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is  
CC also useful in gene therapy. The present sequence is mouse high affinity  
CC choline transporter (mCHT) cDNA. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the US patent office at  
CC [seqdata.uspto.gov/sequence.html?docID=US2004024838](http://seqdata.uspto.gov/sequence.html?docID=US2004024838).

XX Sequence 1743 BP; 406 A; 409 C; 410 G; 518 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	6,05e-314	Length:	1743
Score:	2993.00	Matches:	580
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-10-724-806-4 (1-580) x ADV77918 (1-1743)

QY	1	MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu	20
DB	1	ATGCTTTCCAGTGGAAAGAGCTGGATGCTATATTCCTCTTACCTCTTATTTCTTG	60
QY	21	ValGlyIleTPAlaAlaTyrPlyThrLysAsnSerGlyAsnProGluGluArgSerGlu	40
DB	61	GTTGGAAATATGGGCTGCATGAGAAAACAAAACAGCGGACACCCAGAAAGCGAGTGA	120
QY	41	AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrAla	60
DB	121	GCCATCATATGCGGGGCGGTGACATTTGGTTGTTGGTTTTCATGACAGCC	180
QY	61	ThrTPValGlyGlyGlyTyrTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGly	80
DB	181	ACCTGGGTGGAGAGGCTACATCAATGAGACAGCAAGAGCATGATGGCCAGGTGT	240
QY	81	GlyLeuAlaTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu	100
DB	241	GGTCTAGCTGGGCTCATGSCAACCATTTGATATTCCTGAGCTTAATTTAGTGGTCTG	300
QY	101	PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln	120
DB	301	TTTTTTGGGAAACCTATGGCTTCCAAAGGATATGTGACTATGTATGACCATTCAACAG	360
QY	121	IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe	140
DB	361	ATCTATGGAAGCGATGGGTGGGCTGCTTCATCTCCGCACTGATGGGAGAGATGTC	420
QY	141	TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp	160
DB	421	TGGGCTGACGAAATTTCTCTGCTCATTAAGGGCCACCATCAGCGTATCTTATGAGAT	480
QY	161	ValAlaIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly	180
DB	481	GTTAACATATCGGTATGCTCTGCACTATTCGACATTCCTTATATCCCTAGTGGGTGG	540
QY	181	LeuTyrSerValAlaIleTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr	200
DB	541	CTTACTCTGTGGCAATATCTGATGTGTCCAGCTATTTCTGCAATTTTAAATAGACTGTGG	600
QY	201	IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal	220
DB	601	ATCATGTGCTTTTGGCCCTGTCACTCTGTGACGTACCGACATCGGATTCACAGCTGTG	660

QY	221	HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGlnSerValGluValTyrThrTyrP	240
DB	661	CATGCTAATATACAGAGTCCCTGGCTGGGAACCATTTGAATCATGTTGAGTCACTGAG	720
QY	241	LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrPglAlaTyrPheGlnArg	260
DB	721	CTTGTAATTTTCTGTATTTATGCTGGGTGGAATCCATGGCAAGCCTACTCTCCAGAGG	780
QY	261	ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGly	280
DB	781	GTCCCTCTTCATCTCTGACCAACCATCTCAGAGTACTGCTCTCTGGAGCTTTTGGG	840
QY	281	CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyrP	300
DB	841	TGCCGTGGATGGCTCTACCGCCCATATGATAGAGAGCTATTTGGAGTCTCCACAGACTGG	900
QY	301	AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluAlaAspMetIleLeu	320
DB	901	AACCACTGCTCTACGGGTATCCAGATCCCAAGACTTAAGAGAAAGACATGATTTCTC	960
QY	321	ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal	340
DB	961	CCGATCGTTCGCACTACCTCTGCGCCCTGTATCATCTCTTCTTGGGCTGTGCTGTT	1020
QY	341	SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla	360
DB	1021	TCAGCTGCTGATCTCTCTCAGCTGATCTGTCATCTGTCGCGCAGTTCTATGTTGCT	1080
QY	361	ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal	380
DB	1081	CGGAATATCTACACACTTTCTTCAGACAAATGATCATGACAGAAAGAAATGTGTGGGTC	1140
QY	381	MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys	400
DB	1141	ATGAGATCATCGTCTGTTGTGGAGCATCTGCAACAGCATGCTTGTGTCGACGAAG	1200
QY	401	ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln	420
DB	1201	ACTGTGTATGGGCTCTGTGATCTGACCTGCTGACTCTGTCTATCATATCTTCCACAG	1260
QY	421	LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle	440
DB	1261	CTGCTCTGTATCTTTATCATCAAGAAACCAACCTTATGGGCGCTTGTCTGTTATATT	1320
QY	441	PheGlyLeuPheLeuArgIleThrGlyGlnProTyrLeuTyrLeuGlnProLeuIle	460
DB	1321	TTTGACATATTCTGAGATTTACTGAGAGAGCCATATCTATACCTTGAGCCCTTAATC	1380
QY	461	PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys	480
DB	1381	TTCTACCTCGTTATTTACTCTGACAAAGATGGTATATCAACCAAGGTTCCCATTTTAA	1440
QY	481	ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaIysTyr	500
DB	1441	ACTCTCTCATGATTAATCTCATTTCTTACCAACATTTGTGTTCTTATCTAGCCAAAGT	1500
QY	501	LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg	520
DB	1501	CTATTTGAAAGTGAACCTTGCTCCAAATATAGATGATTTGATGCTGTTGCTCCAGAG	1560
QY	521	HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn	540
DB	1561	CACAGTGAAGAGAAATGAGAACCACTTCTAGTCAGAAAAGAAATATCAAAATTTAAT	1620
QY	541	GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys	560
DB	1621	GAACTTGCACTGTGAAACCTTGCGAGAGCTTAACCTTCAGTTCACTTCCCAATTAAG	1680
QY	561	GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGluAspAsnLeuGln	580
DB	1681	GAGGCCCTCTCTTGAATGATTCCAGTCCGAGGGGGGTCTGGGAGCTGAAGATATTTACA	1740



```
RESULT 5
AAFB1713
ID AAFB1713 standard; cDNA; 1743 BP.
AC AAFB1713;
XX
XX 01-JUN-2001 (first entry)
XX
XX Mouse high affinity choline transporter protein encoding cDNA.
DE
XX
XX high affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
XX ss.
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX FT 1..1743
XX FT CDS /product= "high affinity choline transporter"
XX
XX MO200116315-A1.
XX
XX 08-MAR-2001.
XX
XX 18-AUG-2000; 2000MO-JP005545.
XX
XX 27-AUG-1999; 99JP-00240642.
XX PR 27-DEC-1999; 99JP-00368991.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Haga T, Okuda T;
XX
XX MPI; 2001-226688/23.
XX DR P-PSDB; AAB74666.
XX
XX New rat and human spinal cord high affinity choline transporters, useful
XX PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
XX PT treating Alzheimer's disease.
XX
XX Claim 12; Page 78-82; 90pp; Japanese.
XX
XX The present sequence encodes a mouse (Mus musculus) high affinity choline
XX CC transporter protein designated cho-1. The cho-1 protein has nontropic and
XX CC neuroprotective activities. The cho-1 polynucleotide and protein can be
XX CC used for the diagnosis of diseases related to the expression of cho-1 by
XX CC comparing the cho-1 polynucleotide sequence in a sample to that of a
XX CC control. Drug compositions containing the cho-1 protein or expression
XX CC promoters or inhibitors of cho-1 are useful for treating disorders
XX CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
XX Sequence 1743 BP; 407 A; 410 C; 409 G; 517 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,12e-311 Length: 1743
Score: 2968.00 Matches: 576
Percent Similarity: 99.5% Conservative: 1
Best Local Similarity: 99.3% Mismatches: 3
Query Match: 99.2% Indels: 0
DB: 4 Gaps: 0
US-10-724-806-4 (1-580) x AAFB1713 (1-1743)
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleuPheThrLeuLeuIlePheLeu 20
Db 1 AUTCCTTCCACGTAGAGAGCTGTGCTATTATCTCTTCTTCACTCCCTTAATATTCG 60
QY 21 ValGlyIleTrrPalaIaATrPlySThrLyASnSerGlyASnProGluGluATrSerGlu 40
Db 61 GTTGGAATATGGCTGATGAGAAACCAAAACAGGGGCAACCAAGAGAGCAAGTGA 120
QY 41 AlaIleIleValAlaGlyIleGlyIleuValAlaGlyIlePheThrMetTrpAla 60
|||||
```

```
Db 121 GCCATCATAGTCGGGGCCGTGACATTGGTTGTGGTTGGTTTATCCATGACAGCC 180
QY 61 ThrTrpValGlyGlyIleThrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 181 ACCTGGTTGGAGAGGCTACATCAATGGGACAGCAAGCAAGTGTATGGCCAGTTGT 240
QY 81 GlyLeuAlaITrrPalaHisAlaProIleGlyTyrSerLeuSerLeuIleuGlyIleu 100
Db 241 GGTCTAGCTTGGGCTCAGGACCCATTGGATATTTCTGAGTCTAATTTAGGTGCTG 300
QY 101 PhePheAlaIleProMetIleArgSerIleGlyTyrValThrMetLeuAspProPheLyS 120
Db 301 TTTTTCGCAAACTATGGCTTCAGGAGGATATGACTATGTAAACCATTTTCAACAG 360
QY 121 IleTyrGlyIleArgMetGlyIleuLeuPheIleProAlaLeuMetGlyIleuPhe 140
Db 361 ATCTATGAAAGCGCATGGGTGGGCTGCTCTTCATCTCCCTGCACTGATGGAGAGATTC 420
QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGCAGCAATTTTCTTGCAATTAGGGGCCACCATCAGCGTATCATTTGATG 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 481 GTGACATATCGGTCAATGTCTCTGCACTCATTCATTCCTTTATACCCCTAGGGGTGG 540
QY 181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTrp 200
Db 541 CTCTACTCTGTGGCATATCTGATGTGTGTCAGCTATTCGCAATTTTATAGAGCTGTG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGTCTCCCTTTTGCCCTGTCACTCCGCACTCAGCGACATCGGATTCACAGCTGTG 660
QY 221 HisAlaIleTyrGlnSerProTrpLeuGlyThrIleGluSerValGluValTyrThrTrp 240
Db 661 CATGCTAAATATACAGAGTCCCTGGCGGAGCAACATGAATCAGTTGAAGCTTCACCTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTrpGlnAlaTyrPheGlnArg 260
Db 721 CTGTATATTTTCTGTATATGATGCTGGGGTGAATCCATGGAGGCTATCTCCAGAGG 780
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGluValLeuSerPheLeuAlaAlaPheGly 280
Db 781 GTCTCTCTTCAATCCCAAGCACTATGCTCAGGTACTGTCTTCCGGAGGCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 841 TGCCTGTATGCTTACCCCGCATATGATGAGCTAATTGGAGCTTCCACAGACTGG 900
QY 301 AsnGlnThrAlaITyrGlyTyrProAspProIlySThrLyGluGluAlaAspMetIleu 320
Db 901 AACCAAGATCGCTACGGGTATCCAGATCCCAAGACTAAGAGGAAACAGACATGATTTCTC 960
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCGATGCTTCTGAGTACTCTGCTCCCTGTGTATCATCTCTTGTGGGCTGGTGTCTT 1020
QY 341 SerAlaIleValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetMetPheAla 360
Db 1021 TCAGCTGCTGTCATGCTCCAGCTGACCTGCTCATCTGCGCGAGTTCTATGTTTGGCT 1080
QY 361 ArgAsnIleTyrGlnLeuSerPheAspGlnAsnAlaSerAspLySgIleValITrrVal 380
Db 1081 CGGAAATCTACCAAGCTTTCTTCCACAAATGATCAGACAAAGAAATGTGTGGTCTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaIThrAlaMetAlaLeuLeuThrLyS 400
Db 1141 ATGAGGATCATCTGTGCTTGTGTCGAGCATCTGCAACAGCCATGTGCTTGAGCAAG 1200
QY 401 ThrValTyrGlyLeuTrrPlyTyrLeuSerSerAspLeuValTyrIleIlePheProGln 420
Db 1201 ACTGTATATGGGCTGTGTAACCTGAGCTCTGACCTTGTCTTACATCATCTTCCACAG 1260
|||||
```

OY	421	LeuLeuCyValLeuPheIleIySgIyThrAsnThryrGlyAlaValAlaGlyTyrrIle	440
Db	1261	CTGCTCTGTGACTTCATCATAAAGAAACCAACCTATATGGGGAGATTGGTGGTTAAATT	1320
OY	441	PheGlyLeuPheLeuArgIleThrGlyGlyGluProIyLeuTyrrLeuGlnProLeuIle	460
Db	1321	TTTGGACTATTCCGTAGAGATTACTGGAGGAGAGCCATTCATATCTTGACAGCCCTTATC	1380
OY	461	PheTyrrProGlyTyrrTyrrSerAspIyAsnGlyIleTyrrAsnGlnArgPheProPheIys	480
Db	1381	TTCAACCTGGTTTATCTATCTGCACAAAGATGGTATATACATTCAGAGGTTCCCACTTTAAA	1440
OY	481	ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrrLeuAlaIysTyrr	500
Db	1441	ACTCTCTCCAGGTTCACCTCATCTTTTNCACAACATTGTGTCTTATCTNCCCAAGTAT	1500
OY	501	LeuPheGlySerGlyTyrrLeuProProIyLeuAspValPheAspAlaValAlaIarg	520
Db	1501	CTATTGAAATGGGAACCTGCTCCTCCCAAAATTAGATGATTTGAGTGCCTGTGGCAAG	1560
OY	521	HisSerGlyGluAsnMetAspIySerThrIleLeuValArgAsnGluAsnIleIyLeuAsn	540
Db	1561	CACGTGAAGAAACATGACAAAGCCATTTCTAGTCAGAAATGAAAATCAAAATTAAAT	1620
OY	541	GluLeuAlaProValIyProArgGlnSerIleuThrrLeuSerSerThrPheThrAspIys	560
Db	1621	GAACTTGCACTGTGAAACCTCGGAGAGGCTTAACCTCACTTACCACTTTCACCAATATAG	1680
OY	561	GluAlaLeuLeuAspValAspSerSerProGlyGlySerGlyTyrrGluAspAsnLeuGln	580
Db	1681	GAGGCGCTCCCTTGATGATTCACAGTCCGAGAGGGCTTGGAGCTGAAGATACCTTCACA	1740

of disorders of fuel metabolism such as obesity or diabetes.

Claim 2, Fig 3, 83pp. English.

The present sequence is mouse P4P6B1 cDNA which encodes OMA (obese mice adipocyte) protein. The P4P6B1 fragment was generated by RNA fingerprinting using random primers p4 and p6. OMA is used as a diagnostic reagent for diagnosing a disorder of fuel metabolism in an underweight or an overweight individual, by detecting the transcription level of a gene encoding OMA, which is induced or repressed in an individual by a factor such as genetic obesity, fasting and refeeding of a fasted individual. OMA is useful in the generation of antibodies, for use in pharmaceutical compositions and for studying DNA/protein interactions. Nucleic acids encoding OMA are involved in gene therapy. An inhibitor of OMA or an antisense oligonucleotide that inhibits expression of OMA are useful for treating disorders of fuel metabolism such as obesity or diabetes

Sequence 4938 BP; 1436 A; 1012 C; 976 G; 1514 T; 0 U; 0 Other;

```

QY 221 HisAlaIeTyGlnSerProTyrPleuGlyThrIleGluSerValGluValTyrThrTrp 240
Db 907 CATGCTAAATACCAAGTCCCTGGCTGGGAACCATTAATCACTTGAAGCTTCACTTGG 966
QY 241 LeuAspAsnPhelLeuLeuMetLeuGlyIleProTyrPglAlaTyrPheGlnArg 260
Db 967 CTTGATTAATTTCTGTATTAATGATGCTGGGATGGAATCCATGCAAGCTTAATCCAGAG 1026
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db 1027 GTTCCTCTCTTCATCCCTCAAGCACTATGCTCAGTACTGCTCTTCTGGAGCTTTTGGG 1086
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 1087 TGCTGTGATAGCTTACCCCGCATATGATATGAGCTATTTGGAGCTTCCACAGACTGG 1146
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320
Db 1147 AACCAAGACTGCTACGGGTATCCAGATCCCAAGACTAAGAGAGAGACATGATTTCTC 1206
QY 321 ProIleValLeuGlnThrLeuCysProValTyrIleSerPheGlyLeuGlyAlaVal 340
Db 1207 CCGATGCTTCGACAGTACCTTGCCTGTGTACATCTCTTCTTGGGCTGTGGTGT 1266
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1267 TCAGCTGCTGTATGCTTCCCTCAAGCTGATCTTCATCTGTGGCGAGTTTATGTTTGGT 1326
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal 380
Db 1327 CGGAATATCTACAGCTTTCCTTCCAGCAAAATGATCAGACAGAAATGTGTGGGTC 1386
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1387 ATGAGAGATCACTGTGCTTGTGTTCGAGCATTCGCAACGCCATGTGCTTGCAGAGAA 1446
QY 401 ThrValTyrGlyLeuThrPyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1447 ACTGTATAGGGCTCTGGTACCTGAGCTTCTGACCTTGTCTACATCATCTTCCACAG 1506
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1507 CTGCTCTGTGTACTCTTCAACAAGAGAACCAACATTATGGGCGAGTGTGTTATATT 1566
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyProTyrLeuTyrLeuGlnProLeuIle 460
Db 1567 TTTGGACTATTTCTGGAATTTACTGAGAGAGACCATATCTATATCTTGACGCCCTTAATC 1626
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db 1627 TTCTACCCCTGGTATTAATCTGCAAGAAATGATATACATACAGAGGTCCCATTTAA 1686
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db 1687 ACTCTCTCCATGTTACCTCATCTTTACCAACATTTGTGTTCTTATATCAGCAAGTAT 1746
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
Db 1747 CTAATTGAAAGTGGAACCTTGCTCCAAATTTAGATGATTTGATGTGTGGCAGAGG 1806
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db 1807 CACAGGAAAGAGACATGAGCAAGACCATTTCTAGTCAAGAAATGAAATATCAAAATTAAT 1866
QY 541 GluLeuAlaProValLysProArgIleSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1867 GAACCTTGACACTGTGAAACCTCGGAGAGGCTTAACCTCACTTCAACCAATAAG 1926
QY 561 GluAlaLeuLeuAspValaAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db 1927 GAGGCCCTCTTATGTATGATTCAGTCCGAGAGGGGTCTGGAGCTGAAGATTAATCAAA 1986

```

```

RESULT 7
AAAF81711
ID AAAF81711 standard; cDNA; 1743 BP.
XX
AC AAAF81711;
XX
DT 01-JUN-2001 (first entry)
XX
DE Rat high affinity choline transporter protein encoding cDNA.
XX
KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
XX ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 1..1743
FT FT /*tag= a
FT /product= "high affinity choline transporter"
XX
PN MO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000MO-JP005545.
XX
PR 27-AUG-1999; 99JP-00240642.
PR 27-DEC-1999; 99JP-00368991.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Haga T, Okuda T;
XX
DR WPI; 2001-226688/23.
XX
DR P-PSDB; AAB74664.
XX
PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
PS Claim 6; Page 64-68; 90JP; Japanese.
XX
XX The present sequence encodes a rat (Rattus norvegicus) high affinity
XX choline transporter protein designated cho-1. The cho-1 protein has
XX nootropic and neuroprotective activities. The cho-1 polynucleotide and
XX protein can be used for the diagnosis of diseases related to the
XX expression of cho-1 by comparing the cho-1 polynucleotide sequence in a
XX sample to that of a control. Drug compositions containing the cho-1
XX protein or expression promoters or inhibitors of cho-1 are useful for
XX treating disorders characterised by abnormal levels of cho-1, such as
XX Alzheimer's disease
XX
SQ Sequence 1743 BP; 414 A; 402 C; 404 G; 523 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6,266-310 Length: 1743
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 98.8% Indels: 0
DB: Gaps: 0
US-10-724-806-4 (1-580) x AAFA81711 (1-1743)
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGGCTTTCCATGTATAGAGACTAGTACGATTAATCTGTCTTACCTTCTTATATTTCTG 60
QY 21 ValGlyIleTyrAlaIleTyrPylsThrLysAsnSerGlyAsnProGluGluArgSerGlu 40
Db 61 GTTGGAAATATGGCTCATAGGAAACCAAAACAGCGGTATATCCAGAAACGACGAA 120
QY 41 AlaIleIleValaIglyIlyArgAspIleGlyLeuLeuValaIglyIlyPheThrMetThrAla 60

```

```

Db      |||||
121  GCCATCATAGTGGGGCCGAGACATTGGTTGGTGGTGGTTTCCATGACAGCC 180
Qy      ThrTPValIglYglYTyrlleasnGlyThrAlaIglValYlYrGlyProGlyCys 80
Db      181  ACCTGGGTGGAGAGGTTACATCAACGGGACAGCTGAAGCACTTATGGCCAGGTTGT 240
Qy      GlyLeuAlaTPAlaHisAlaProIleGlyTySerLeuSerIleuIleuGlyGlyLeu 100
Db      241  GGCTAGCTGGGGCTCAGGACCCCATTTGGATATCTCTGAGTCTGATTTTAGGGCTCG 300
Qy      PhePheAlaIysProMetArgSerIysGlyTyTrValThrMetLeuAspProPheIysGln 120
Db      301  TTTTTCGAAAACCTATGCGTTCCAGGGATATGACTATGTTAACCCTGGTTCAACG 360
Qy      IleTyrglyLysArgMetGlyGlyLeuLeuPheIleProAlaIleuMetGlyGlnMetPhe 140
Db      361  ATCTATGAAACCGCATGGGTGGGCTGCTGTTCACTCCGCACTGATGGGAGAGATGTTG 420
Qy      TrpAlaAlaAlaIlePheSerAlaIleuGlyAlaThrIleSerValIleIleAspValAsp 160
Db      421  TGGGCTGCAGCAATTTCTCTCATATGAGGGCTACCATCAGCGTATCATGATGAGAT 480
Qy      ValaenIleSerValIleValSerAlaIleuIleAlaIleuTyThrIleuValIglYglY 180
Db      481  GTGAACATATCGGTCAATGTCCTCCGCACTCATTTGCCATTTCTTATACCTCGTGGAGGG 540
Qy      LeuTySerValAlaIATyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      541  CTCTACTCTGTGGCATATCTGATGTGTACAGCTATTCGCACTTTTATAGATTTGTGG 600
Qy      IleSerValProPheAlaIleuSerHisProAlaValThraspIleGlyPheThrAlaVal 220
Db      601  ATCAATGTCCTCAATGGCCCTGTCACATCTCGCAGTCACGACCATTTGATTCACCTGCTGG 660
Qy      HisAlaIysTyrglySerProTrpIleuGlyTyrlleGlnSerValGlnValTyThrTrp 240
Db      661  CATGCTAAATACCAAGTCTCCGCTGGGAAACCATTTGAATCGATGAATCTACACCTGG 720
Qy      LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaIATyrPheGlnArg 260
Db      721  CTGTATATTTTCTGTGTGTGATGCTGGGTGGAATTCATGGCAAGCTACTTCCAGAG 780
Qy      ValLeuSerSerSerSerAlaThrTyrlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      781  GTCCCTCTCTTCATCGTCACGACCTATGCTCAGGTGCTCTCTCTCTGACAGCTTTGGG 840
Qy      CysLeuValMetAlaIleuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db      841  TGCTGTGATGAGCTTACCAAGCATTTGAGGAGGCAATGGACCTCCACAGACTGG 900
Qy      AsnGlnThrAlaIATyrGlyTyTrProAspProIySThrIysGlnGlnAlaAspMetIleLeu 320
Db      901  AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGAGAGACATGATTC 960
Qy      ProIleValLeuGlnTyrlleuGlyCysProValTyrlleSerPhePheGlyLeuGlyAlaVal 340
Db      961  CCGATGTTCTTACAGTACCTCTGCTGTGTACATTTCTTCTTGGGCTGTGCTGT 1020
Qy      SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db      1021  TCTGCTGCTGTATGCTCTCGGCTGACTGATCTCATCTTATCCAGCAAGTCCATGTTGCT 1080
Qy      ArgAsnIleTyrglyLeuSerPheArgGlnAsnAlaSerAspIysGlnIleValTrpVal 380
Db      1081  CGGAATATCTACACAGCTTCTCTTCCAGACAAATGCATCAGACAAAGAAATTTGTGGGCTC 1140
Qy      MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaIleuLeuTrpIys 400
Db      1141  ATGAGATCATACGTGTGTGTGTGGAGCATCTGCACAAAGCATGGCTTGTCTCAGGAAG 1200
Qy      ThrValTyrglyLeuTrpTyrlleuSerSerAspLeuValTyrlleIleIlePheProGln 420

```

```

Db      1201  ACTGTGATAGGGCTCTGTGATACCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260
Qy      LeuLeuCysValIleuPheIleYsGlyThrAsnTrpTyrglyAlaValAlaGlyTyrlle 440
Db      1261  CTGCTCTGTGTACTCTTTCATCAAAAGAACCAACACTTATGGGCGAGTCTGCTATAT 1320
Qy      PheGlyLeuPheLeuArgIleThrGlyGlyIleProTyrlleuTyrlleuGlnProLeuIle 460
Db      1321  TTTGACATTTTCTCAGATTAACCGAGAGAGAGCCATATCTATCTTGACAGCCCTTAATC 1380
Qy      PheTyTrProGlyTyTrTySerAspIysAsnGlyIleTyAsnGlnArgPheProPheIys 480
Db      1381  TTCTACCTGCTTATTAACCTGACAAAGGTATATACATCAAGAGGTTCCCATTTAA 1440
Qy      ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrlleuAlaIysTy 500
Db      1441  ACTCTCTCATATGTTACCTCATCTTTTCCAAACATTTTGTGTTCTCATCTACCCAAAT 1500
Qy      LeuPheGlnSerGlyThrLeuProProIysLeuAspValPheAspAlaValAlaArg 520
Db      1501  CTATTTGAAAGTGAACCTTGCTCCAAAAATTAGATATATTTGATGAGCTGTGCTCAAG 1560
Qy      HisSerGlnGlnAsnMetAspIysThrIleLeuValArgAsnGlnAsnIleIysLeuAsn 540
Db      1561  CACAGTGAAGAGACATGAGACAGACCATTTCTATGTCAGAAATGAAACATCAAAATTAAT 1620
Qy      GluLeuAlaProValIysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIys 560
Db      1621  GAACTTGACCTGTAAAGCTTGACAGAGCCTAACCTCAGTTCACATTTCCCAATATA 1680
Qy      GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyTyThrGlnAspAsnLeuGln 580
Db      1681  GAGGCTCTCCTTGAATGATTCAGTCCAGAGGAGATCGGACATGAAGATTAATTACAA 1740

```

RESULT 8

ADD50642

ID ADD50642 strand; cDNA; 4904 BP.

XX

AC ADD50642;

DT 15-JAN-2004 (first entry)

XX

DE cDNA encoding rat high-affinity choline transporter (rCHT).

XX

KW Rat; high-affinity choline transporter; rCHT; cholinergic function;

KW Parkinson's disease; Huntington's disease; Alzheimer's disease;

KW schizophrenia; dysautonomia; myasthenia gravis; brain;

KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;

KW neuroprotective; neuroleptic; gene; ss.

XX

OS Rattus sp.

XX

XX

FX Key Location/Qualifiers

FT CDS 224..1966

FT /\*tag= a

FT /product= "rCHT"

PN

XX

US2003114399-A1.

PD 19-JUN-2003.

XX

PF 23-JUL-2001; 2001US-00911077.

XX

PR 23-JUL-2001; 2001US-00911077.

XX

PA (BLAK/) BLAKELY R D.

PA (APPA/) APPARUNDARAM S.

PA (FERG/) FERGUSON S.

XX

PI Blakely RD, Apparsundaram S, Ferguson S;

XX

DR WPI: 2003-810914/76.

DR P-PSDB; ADD50643.

XX Novel isolated polynucleotide encoding human or mouse high affinity  
PT choline transporter polypeptide, useful in gene therapy to increase  
PT cholinergic function in a cell of a patient suffering from Alzheimer's  
PT disease.

XX Example 1; SEQ ID NO 5; 74bp; English.

XX The present invention relates to the isolation of polynucleotide  
CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic  
CC function in the cell that is in a patient having Parkinson's disease  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signalling. The present sequence encodes rat CHT (rCHT). Note: The  
CC sequence data for this patent was obtained in electronic format directly  
CC from the USPTO web site at seqdata.uspto.gov.

XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

Alignment Scores:

Score: 3.06e-309 Length: 4904  
Percent Similarity: 2956.00 Matches: 572  
Best Local Similarity: 99.3% Conservative: 4  
Query Match: 98.6% Mismatches: 4  
10 Indels: 0  
Gaps: 0

US-10-724-806-4 (1-580) x ADD50642 (1-4904)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
DB 224 ATGCTTTCATGTAAGAGCATGTAAGCATTAATCTGTTCACTTCTTAATTTCTG 283  
QY 21 ValGlyIleTyrAlaIleTyrIleThrIleAsnSerGlyAsnProGluIleArgSerGlu 40  
DB 284 GTTGGAAATATGGGCTGCATGAGAAACCAAAACAGCGGTATGCAGAGAACAGCGAA 343  
QY 41 AlaIleIleValGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60  
DB 344 GGCATATATGTTGGGGCCGAGACATTTGTTGTTGGTGTGTTTACCATGAGACGCC 403  
QY 61 ThrTyrValGlyIleGlyIleThrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
DB 404 ACCTGGGTTGGAGAGGAGTTACATCAACGGACAGCTGAGACGATTATATGGCCAGTTGT 463  
QY 81 GlyLeuAlaIleTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyIleLeu 100  
DB 464 GGTTCTAGCTTGGGCTCAGGCACCCATTGGATATTCTCTGAGTCTGATTTAGGTGGCTG 523  
QY 101 PhePheAlaIlePheProMetIleArgSerIleGlyTyrValThrMetLeuAspProPheLeu 120  
DB 524 TTTTGTGAAAACCTATGCGTTCCAAAGGATATGACTATGTATGACCCGTTTCAACG 583  
QY 121 IleTyrGlyIleArgMetGlyIleLeuPheIleProAlaLeuMetGlyIleMetPhe 140  
DB 584 ATCTATGAAAGAGGAGTGGGCTGGGCTGCTTTCATCCCTGCACATGATGGAGAGATGTT 643  
QY 141 ThrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
DB 644 TGGGCTGCAGCAATTTCTCTGCTATTAAGGGCTTACCATCAGCCCTAATCATTTGATGGAT 703  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyIle 180  
DB 704 GTGAACATATCGGTCTTGTCTCCGACATCATTTGCAATTTTATACCTTCGAGGAGGG 763  
QY 181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTyr 200

DB 764 CTTACTCTGTGGCATATCTGATGTTGTATACAGCTATTCGATTTTATAGATTTGG 823  
QY 201 IlesSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyIlePheThrAlaVal 220  
DB 824 ATCAGGTCCCATTTGGCTGCTGCATCCTGCAGACCGAATGATGATCTGCTG 883  
QY 221 HisAlaIleTyrGlnSerProTyrLeuGlyIleIleGlySerValGluValTyrThrTyr 240  
DB 884 CATGCTAATATACAGAGTCCCTGCTGGAGACCATTTGATGATGATGATGATGAT 943  
QY 241 LeuAspAsnPhenLeuLeuLeuMetLeuGlyIleProTyrGlnAlaTyrPheGlnArg 260  
DB 944 CTGTAATATTTCTGTTGTTGATGCTGGGGTGAATACATGCGAACCTTATCTCCAGAG 1003  
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
DB 1004 GTCTCTCTTCATCGTCAGACGACTATGCTCAGGTCTGCTCTGCGAGCTTTTGG 1063  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
DB 1064 TGCTGTGATATGCTTACAGCATTTGCTGAGGCTGATGGAGCTTCCAGACTGG 1123  
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrIleGlyGluAlaAspMetIleLeu 320  
DB 1124 AACCAATGCAATATGGGTTTCCAGATCCCAAGACCAAGAGAGACAGCATGATTTCT 1183  
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
DB 1184 CCGATTGTTCTACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
DB 1244 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303  
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIleGlyIleValIleTyrVal 380  
DB 1304 CGGAATATCTACAGCTTTCTTCCAGAAATGATGATGATGATGATGATGATGATGAT 1363  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400  
DB 1364 ATGAGGATCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1423  
QY 401 ThrValTyrGlyIleLeuTyrIleLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
DB 1424 ACTGTATAGGGCTCGTACCTGAGCTGAGCTTGTCTACATCATCTTCCACAG 1483  
QY 421 LeuLeuCysValLeuPheIleIleIleGlyIleThrAsnThrTyrGlyAlaValAlaGlyIle 440  
DB 1484 CTGCTCTGTACTCTTCTTCAAGAGAACCAACACTTATGGGCACTGCTGTATATT 1543  
QY 441 PheGlyLeuPheLeuArgIleThrGlyIleGluProTyrLeuTyrIleGlnProLeuIle 460  
DB 1544 TTTGACTTTTCTGTAATTAACCGAGAGAGCCATCTAATCTAATCTGAGCCCTTAATC 1603  
QY 461 PheTyrProGlyTyrTyrSerAspIleAsnGlyIleTyrAsnGlnArgPheProPheLeu 480  
DB 1604 TTCTACCTGCTTATTAACCTGACAGAAATGATATACATATAGAGGTTCCATTTAA 1663  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500  
DB 1664 ACTCTCTCATGGTATACCTCATCTTATCAACATTTGTTGTTCTATGACCAAGTAT 1723  
QY 501 LeuPheGlnSerGlyIleThrLeuProIleLeuAspValPheAspAlaValAlaIleArg 520  
DB 1724 CTATTTGAAAGTGAACCTTGGCTCCAAATTAATTAATTTGATGCTGTTGCTCAAG 1783  
QY 521 HisSerGlyIleAsnMetAspIleThrIleLeuValArgAsnGluAsnIleIleLeuAsn 540  
DB 1784 CACAGTGAAGAGACATGAGCAACATCTTCTAGTCAAGAAATGAAATCAATCAATTAAT 1843  
QY 541 GluLeuAlaProValIleProArgIleSerLeuThrLeuSerSerThrPheThrAsnIle 560

Db 1844 GAACCTGCACCTGTAAGCCTGCAGAGCCTACCTCAGTTCACTTACCAATATA 1903

QY 561 Glu1AlaLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
 |||||  
 Db 1904 GAGGCTCTCCCTTGATGATTGATTCAGAGGAGATCTGGGACATGAGATACCTTACA 1963

RESULT 9

ID ADV77900 standard; cDNA; 4904 BP.

XX ADV77900;

AC ADV77900;

XX 24-FEB-2005 (first entry)

DT Rat high affinity choline transporter (rCHT1) cDNA.

DE

XX Choline transporter; neuromuscular disorder;

KM autonomic nervous system disorder; central nervous system disorder;

KM Parkinsons disease; Huntingtons chorea; genetic disorder;

KM Alzheimers disease; degeneration; neurological disorder; schizophrenia;

KM psychiatric disease; myasthenia gravis; immune disorder; gene therapy;

KM cns-gen.; neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;

XX muscular-gen.; neuroleptic; DNA purification; gene; ss.

OS Rattus norvegicus.

XX

XX Key Location/Qualifiers

FT CDS 224.1966

FT /tag= a

FT /product= "Rat CHT protein"

XX

XX US2004248838-A1.

XX

XX 09-DEC-2004.

XX

XX 01-DEC-2003; 2003US-00724806.

XX

XX 23-JUL-2001; 2001US-00911077.

XX

XX (UYVA-) UNIV VANDERBILT.

XX

XX Blakely RD, Apparsundaram S, Ferguson S;

PI WPI; 2005-020586/02.

XX

XX P-PSDB; ADV77901.

DR

XX GENBANK; AB030947.

XX

XX New human high affinity choline transporter (CHT) cDNA, useful for

PT treating a neuromuscular, autonomic or central nervous system disorder,

PT including Parkinson's disease, Huntington's disease, Alzheimer's,

PT schizophrenia.

XX

XX Example 1; SEQ ID NO 5; 73bp; English.

XX

XX The present invention provides polynucleotides encoding novel high

CC affinity choline transporters (CHTs), methods for their use in screening

CC and therapy. The invention is useful for treating neuromuscular

CC disorders, autonomic or central nervous system disorders such as

CC Parkinsons disease, Huntingtons disease, Alzheimers disease,

CC schizophrenia, dysautonomia and myasthenia gravis (MG). The invention is

CC also useful in gene therapy. The present sequence is rat high affinity

CC choline transporter (rCHT1) cDNA. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the US patent office at

CC segdata.uspto.gov/sequence.html?DocID=US20040248838.

XX

XX Sequence 4904 BP; 1447 A; 991 C; 939 G; 1527 T; 0 U; 0 Other;

SO

Query Match: 98.88 Indels: 0

DB: 14 Gaps: 0

US-10-724-806-4 (1-580) x ADV77900 (1-4904)

QY 1 MetProPheH1sValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
 |||||

Db 224 ATGCTTTCATGTAAGAGAGACTAGTACGATATCTGTTTACCTTATATTTCTG 283

QY 21 ValGlyIleTrrPalaaIaTrrPlySThrLySaSnsSerGlyAsnProGluGluArgSerGlu 40  
 |||||

Db 284 GTTGGAAATATGGGCTGCATGGAACCAAAAACAGCGGTAAATGCAGAGAACGACGCA 343

QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
 |||||

Db 344 GCCATCATAGTGGGGCCGAGACATTTGTTGTGGTGGGTTTAAACATGACAGCC 403

QY 61 ThrTrrPalGlyGlyGlyTrrIleSngIlyThrAlaGluAlaValTyrGlyProGlyCys 80  
 |||||

Db 404 ACCTGGTTGGAGGGGTTTACATCAACGGAGAGCTGAAGAGATTATAGGCCAGAGTTGT 463

QY 81 GlyLeuAlaTrrPalaaIleAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
 |||||

Db 464 GGCTAGCTGGGCTCAGAGCAACCATGGATATCTCGAGTCTGATTTAGTGCCCTG 523

QY 101 PhePheAlaIlySPromerArgSerIlyGlyTrrValTrrMetLeuAspProPheIlySgln 120  
 |||||

Db 524 TTTTTCGAAAACCTATGCGTTCCAGAGGATATGTGACTATGTATGACCCTTTCAACAG 583

QY 121 IleTrrGlyLySaArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuMetPhe 140  
 |||||

Db 584 ATCTATGGAAGCGATAGGTGGCTGCTTTCATCCCTGCACATATGGAGAGATGTTTC 643

QY 141 TrrPalaaIaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValIle 160  
 |||||

Db 644 TGGGCTGCAGCAATTTCTCTGCATTAGGGCTACATCAGGATCATGATCATGTGAT 703

QY 161 ValaenIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
 |||||

Db 704 GTGAACATATGCGTCAATGTCCTCGCACTCATTCGCAATTTTATACCTCTGGAGGG 763

QY 181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrr 200  
 |||||

Db 764 CTCTACTCTGGGCAATATACATGATTTGTACAGCTATTTCTGATTTTATAGGATTTGG 823

QY 201 IleSerValProPheAlaLeuSerHisProAlaValTrrAspIleGlyPheThrAlaVal 220  
 |||||

Db 824 ATCAGTGTCCCATTTGGCCTGTCAATCTGCAGTCCAGCATTTGATTCTGCTG 883

QY 221 HisAlaIlyTrrGlnSerProTrrLeuGlyThrIleGlySerValGluValTyrThrTrr 240  
 |||||

Db 884 CATGCTAAATACAGAGTCCCTGGCTGGGAACCATGTGAATCATGTAAGTCTACACCTGG 943

QY 241 LeuAspAnPheLeuLeuLeuMetLeuGlyIleProTrrGlnAlaTyrPheGlnArg 260  
 |||||

Db 944 CTTGATATTTTCTGTTGTGTGATGCTGGGTGAATACCATGGCAAGCTACTTCCAGAGG 1003

QY 261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
 |||||

Db 1004 GTCCCTCTTTCATCTGTCAGCAGCATATCTCAGGGCTGTCTCTCTGAGCTTTTGGG 1063

QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300  
 |||||

Db 1064 TCCCTGGTGAATGGCTCTACAGCCCATTTGGCATTTGGGGCCATTTGGAGCTTCCACAAC 1123

QY 301 AsnGlnThrAlaTyrGlyTrrProAspProIlySThrLySngIlyGluAlaAspMetIleLeu 320  
 |||||

Db 1124 AACCAAACTGCATATGGGTTTCCAGATCCCAAGACCAAGAGAGACATGATTTCTC 1183

QY 321 ProIleValLeuGlnTrrLeuCysProValTrrIleSerPhePheGlyLeuGlyAlaVal 340  
 |||||

Db 1184 CCGATGTTTTCACAGTACCTGACCTGTGATCATTTCTTCTTGGGCTTGCTGTT 1243

```
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1244 TCTGCTGCTGATGATGCTCGGCTGATCATCTCATCTATCAGCAAGTTCCATGTTGCT 1303
QY 361 ArgAsnIleTyrGluLeuSerPheArgGluAsnAlaSerAspGluIleValTyrPAl 380
DB 1304 CGGAATATCTACACCTTCTCCAGACAAAATGATCAGACAGAGAAATTTGTGGGTC 1363
QY 381 MetArgIleThrValLeuValPheGluAlaSerAlaThrAlaMetAlaLeuThrLys 400
DB 1364 ATGAGATCATCTGTGTTGTTGTTGGAGATCTGCACAGCCATGGCTTGTCTCAGAG 1423
QY 401 ThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1424 ACTGTGATGGGCTCTGGTACTGAGCTGACCTGTGATCATCATCATCTTCCACAG 1483
QY 421 LeuLeuCyValLeuPheIleIleGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1484 CTGCTCTGTGATCTTTCATCAAGAACCAACACTTATGGGGCACTTCTGTATATAT 1543
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGlyProTyrLeuTyrLeuGlnProLeuIle 460
DB 1544 TTGGACTTTCTCTGGAATTCAGAGAGAGCCATATCTTACTTGCAGCCCTTAATC 1603
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
DB 1604 TTCTACCCGTGTTATTAACCTGACAAAGATGATATATACATCAGAGGTTCCCATTTAA 1663
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCyValSerTyrLeuAlaLysTyr 500
DB 1664 ACTCTCTCCATGTTACCTCATCTTTTACCAACATTTGGTTCCATCTAGCAAGATAT 1723
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
DB 1724 CTATTTGAAAGTGAACCTTGCTCCAAATTAATGATATATTTGATGCTGTGTCTAAG 1783
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValAlaArgAsnGluAsnIleLysLeuAsn 540
DB 1784 CACAGTGAAGAGAACATGACAAAGACCATTTAGTCAGAAATCAAAACATTAATTAAT 1843
QY 541 GluLeuAlaProValIysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1844 GAACCTTGACCTGTAAAGCCTGCAGACAGACCTTAACCTCAGTTCACTTCCAAATAA 1903
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB 1904 GAGGCTCTCCTGATGATGATTCACACTCCAGAGGGATCTGGGACTGAAAGATTAAC 1963
RESULT 10
AAF81712
ID AAF81712 standard; cDNA, 1743 BP.
XX
AC AAF81712;
XX
DT 01-JUN-2001 (first entry).
XX
DE Human high affinity choline transporter protein encoding cDNA.
XX
KW High affinity choline transporter; cho-1; Alzheimer's disease; diagnosis;
KM ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1743
FT /tag= a
FT /product= "high affinity choline transporter"
XX
PN MO200116315-A1.
XX
PD 08-MAR-2001.
XX
PF 18-AUG-2000; 2000MO-JP005545.
```

```
XX
PR 27-AUG-1999; 99UP-00240642.
PR 27-DEC-1999; 99UP-00368991.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PI Haga T, Okuda T;
XX
DR WPI; 2001-226688/23.
DR P-PSDB; AAB74665.
XX
PT New rat and human spinal cord high affinity choline transporters, useful
PT in diagnosis of Alzheimer's disease and screening promoters as drugs for
PT treating Alzheimer's disease.
XX
PS Claim 9; Page 71-75; 90pp; Japanese.
XX
CC The present sequence encodes a human (Homo sapiens) high affinity choline
CC transporter protein designated cho-1. The cho-1 protein has nootropic and
CC neuroprotective activities. The cho-1 polynucleotide and protein can be
CC used for the diagnosis of diseases related to the expression of cho-1 by
CC comparing the cho-1 polynucleotide sequence in a sample to that of a
CC control. Drug compositions containing the cho-1 protein or expression
CC promoters or inhibitors of cho-1 are useful for treating disorders
CC characterised by abnormal levels of cho-1, such as Alzheimer's disease
XX
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,84e-292 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: Gaps: 0
US-10-724-806-4 (1-580) x AAF81712 (1-1743)
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB 1 ATGGCTTTCATGATGAGAAAGACTGATGATCATCATGTTTCTTACCTTTAATTTGCTG 60
QY 21 ValGlyIleTyrAlaAlaTyrPyrThrLysAsnSerGlyAsnProGluGluArgSerGlu 40
DB 61 GTTGAATATGGGCTGCCCTCGAGAAACCAAAAACAGTGGCAGCGCACAAGCGCACGAA 120
QY 41 AlaIleIleValAlaGlyIleArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCCATCATGTTGGTGGCCGAGATATGTTGTTATTTGTTGGATTTTACATGACAGCT 180
QY 61 ThrTyrPAlaGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB 181 ACCGTGGGTGGAGAGGGTATATCATATGACAGCTGAAAGCAGTTTATGATACAGTTTAT 240
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGCTTACGCTGGGCTCAGCACCAATGATATCTTCTTATCTGCTATTTAGGTGGCTG 300
QY 101 PhePheAlaLysProMetArgSerIysGlyTyrValThrMetLeuAspProPheLysGln 120
DB 301 TTCTTGCAAAACCTATGCGTTCAAAAGGGGTATGTACCATGTTAGACCGTTTCAGCAA 360
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB 361 ATCTATGAAAGACATGGGCGGACTCTGTTTATTTCTGCACTATGGAGAAATGTTCT 420
QY 141 TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGAGCAATTTCTCTGCTTGGAGCCACCATCAGCGTATCATCATGTGGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB 481 ATGACATTTCTGATCATCTCTGCACTCATTTGCCACTGTGACACACTGGTGGAGGG 540
```

```

QY 181 LeuTYrSerValAlaTYrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db 541 CTCAATTCGTGGCCCTGACATGATGTCGTGAGCTCTTTGGCATTTTGTAGGGCGCTGG 600
QY 201 IleserValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGCGTCCCTTTGGATGTCACATCTCGAGTCGAGACATGGGGTTCCTGCTG 660
QY 221 HisAlaIysTYrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTYrThrTrp 240
Db 661 CATGCCAATAACCAAAAGCCGTGGGAACTGTGTACTCATCGAAGTCTACTCTTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTYrPheGlnArg 260
Db 721 CTGTAGATGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 261 ValLeuSerSerSerSerSerAlaThrTYrAlaGlnValIleLeuSerPheLeuAlaIlePheGly 280
Db 781 GTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db 841 TGCCGTGTGATGGCCATCCAGCCATCATTTGGGGCCATTGGAGCATCAACAGCTGG 900
QY 301 AsnGlnThrAlaTYrGlyTYrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320
Db 901 AACCCAGACTGCATATGGGCTTCCAGATCCCAAGACTACAGAAAGGCAACATCATTTTA 960
QY 321 ProIleValIleGlnThrLeuCysProValTYrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATGTTCTGACAGATCTGCCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCTGCTGCTGTTATGTCATCAGACATTTCTTCATCTTGTGCAAGATTCATGTTTGA 1080
QY 361 ArgAsnIleTYrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTYrVal 380
Db 1081 CGGAACATCTACACACCTTCTTCCAGACAAATGCTTCGGAACAAAGAAATCGTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGCGAATCAGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
QY 401 ThrValTYrGlyLeuTrpTYrLeuSerSerAspLeuValTYrIleIleIlePheProGln 420
Db 1201 ACTGTGTAATGGGCTCTGGTACTCAGTCTGACCTTGTTTACATGTTATCTTCCCCAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTYrGlyAlaValAlaGlyTYrIle 440
Db 1261 CTGCTTTGTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTYrLeuTYrLeuGlnProLeuIle 460
Db 1321 TCTGGCCTCTTCTCTAGATTAACGAGGGAGCCATATCTGTATCTTCAAGCCTTGATC 1380
QY 461 PheTYrProGlyTYrTYrSerAspLysAsnGlyIleTYrAsnGlnArgPheProPheLys 480
Db 1381 TTCTACCTGCTATTAACCTGATGATATGATATATATATATATATATATATATATATATA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTYrLeuAlaIysTYr 500
Db 1441 ACACCTGGCAGTATCATCTTCAACCAACATTTGATCTCTCATCTAGCCAAAGTAT 1500
QY 501 LeuPheGlnSerGlyTYrThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
Db 1501 CTATTGAAAGGAACTTGCACCTTAATTAAGATGATTTGATGCTGTGTGTGAAAGA 1560
QY 521 HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db 1561 CACAGCTGAAGAAATATGATATAGACAAATCTTGTCAAAAATGAAAATATTAATTAGAT 1620

```

```

QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGACCTTGTGAACCCAGACAGACATGACCTCAGCTCAACTTTCACCAATATA 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyTYrGluAspAsnLeuGln 580
Db 1681 GAGGCTTCCTTGTATGTTGATTCCAGTCCAGAAAGGCTGTGGACATGAAGATTAATTACAG 1740

RESULT 11
AAH49207
ID AAH49207 standard; cDNA; 1743 BP.
AC AAH49207;
XX 26-NOV-2001 (first entry)
DT XX
XX XX
DE Human CHOT encoding cDNA.
KW CHOT; human; choline transporter; chromosome 2q11-13; nootropic;
KW neuroprotective; gene therapy; antisense therapy; degenerative disease;
KW cognitive disorder; Alzheimer's disease; ss.
OS Homo sapiens.
PN DE1009055-A1.
PD XX
XX 30-AUG-2001.
PF 28-FEB-2000; 2000DE-01009055.
PR 28-FEB-2000; 2000DE-01009055.
PA (BRUE/) BRUESS M.
PA (BOEN/) BOENISCH H.
PI Bruesse M, Boenisch H;
XX
XX WPI; 2001-590709/67.
DR P-PSDB; AAB86837.
XX
PT A new gene encoding human choline transporter, designated hCHOT is
PT located on chromosome 2q11-13 and is useful to treat degenerative
PT disorders such as Alzheimer's disease.
PS
XX Disclosure; Page 11; 12pp; German.
XX
XX This invention describes a novel gene encoding human choline transporter,
XX designated hCHOT which is located on chromosome 2q11-13. The products of
XX the invention have nootropic and neuroprotective activity and can be used
XX for gene or antisense therapy. (1) is used to treat degenerative disease,
XX particularly cognitive disorders such as Alzheimer's disease. Sense and
XX antisense oligonucleotides derived from the gene may be used in
XX CC diagnostics and other techniques. This sequence encodes the human CHOT
XX protein described in the invention
XX
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,84e-292 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: 5 Gaps: 0

US-10-724-806-4 (1-580) x AAH49207 (1-1743)
QY 1 MetProPheHisValGlnGlyLeuValAlaIleIleLeuPheTYrLeuLeuIlePheLeu 20
Db 1 ATGGCTTCCATGCTGGAAGGAGCTAGTACATCATCGTGTCTTACTTATTTTGTCTG 60
QY 21 ValGlyIleTPAlaAlaIleTrpLysThrLysAsnSerGlyAsnProGlnGluArgSerGlu 40

```



Db 61 GTTGAATATGGGCTGCTGAGAACCAAAACAGTGCAGCGCAGAGAGCGCAGCGAA 120  
 QY 41 AIAIIeIIeIValIGlyIArgAspIIeGIyleuLeuValIGlyIphMetThrAla 60  
 Db 121 GCCATATATGGTGGCCGAGATATGGTTATTTGGTGGATTTACCAAGACAGCT 180  
 QY 61 ThrTPValIGlyIYrIIeaenGIYThraIGluAlaValYrGIYProGIYCyS 80  
 Db 181 ACCTGGGTGGAGAGGATATCATGGCAGCGTGAAGCAGTTATATGACAGGTTAT 240  
 QY 81 GlyLeuAlaTPAlaHisAlaProIIeGIYrSerIeuSerIeuIIeGIYIleu 100  
 Db 241 GGCCTTACCTGGGCTCAGGACCAATTTGATATCTTGAATTTAGGTCGCTG 300  
 QY 101 PhePheAlaIleProMetArgSerIYrGIYrValThMetIeuSerProPheIYSGIn 120  
 Db 301 TTCTTTGCAAAACCTATGGCTCAAAAGGGATATGATACATGTTAAGCCGTTGAGCA 360  
 QY 121 IleYrGIYIYsArgMetGIYGIYLeuLeuPheIIeProAlaIeuMetGIYIuMetPhe 140  
 Db 361 ATCTATGAAAAGCATGGGCGGACCTCGTTTATTCCTGCACATGATGGAGAAATGTT 420  
 QY 141 TPAlaAlaIaIaIePheSerAlaIeGIYAlaIaThrIleSerValIIeIleAspValaAsp 160  
 Db 421 TGGGCTGCAGCAATTTCTCTGCTTGGGAGCCACATCAGCGTGATCATCGATGTGAT 480  
 QY 161 ValAsnIIeSerValIIeValSerAlaIeAlaIeAlaIeIeuYrThrIeuValaIGlyI 180  
 Db 481 ATGCACATTTCTGTCAATCTGTGACATCATTCATTCATTCATTCATTCATTCATTCAT 540  
 QY 181 LeuYrSerValaIaIaYrThraSPValaIGluLeuPheCySIIePheIIeGIYLeuYr 200  
 Db 541 CTCTATTCGTGGCCCTACACTGATGTCGTTGACCTCTTTGGCAATTTTGAGGGCTGTGG 600  
 QY 201 IISerValIProPheAlaIeIeuSerHisProAlaValaIThraSPIIeGIYpHeThraIaVala 220  
 Db 601 ATCAGGCTCCCTTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 660  
 QY 221 HIsAlaIeYrGIYIeuSerProTProIleuGIYIleIleGIYSerValaIGluValaYrThr 240  
 Db 661 CATGCCAAATACCAAAAGCGGTGGGAGAACTGTGACCTCAATCAATCAATCAATCAAT 720  
 QY 241 LeuAspAsnIleuIeuIeuMetIeuGIYIleIleProTProIleuAlaYrPheGInArg 260  
 Db 721 CTTGATAGTTTCTGTGTTGATGCTGGGTGGAAATCCATGAGAGCAATCTTTCAGAGG 780  
 QY 261 ValIeuSerSerSerSerAlaThrYrIaIGluValIeuSerPheIeuAlaIaIePheGI 280  
 Db 781 GTTCTCTCTTCTTCTCAGGCACTATGCTCAAGTGTCTGTCTCTGCGAGCTTTCGGG 840  
 QY 281 CysIeuValMetAlaIeuProAlaIeCysIIeGIYAlaIeGIYIaIeSerThraSPTr 300  
 Db 841 TGGCTGGTATGGCCATCCGAGCATATCTATTGGGGCCATTTGGACATCAACAGACTGG 900  
 QY 301 AsnGIYThraIaIaYrGIYrProAspProIYsThIeGIYIuAlaIaSPMetIIeIeu 320  
 Db 901 AACCGACTGCAATATGGCTTCCAGATCCAGACATACAGAGAGGCGACAGATGATTTT 960  
 QY 321 ProIIeValIeuGIYrIeuCysProValYrIIeSerPhePheGIYIeuGIYAlaVala 340  
 Db 961 CCNAATGTTCTGAGATCTCTGCGCTGATATATTTCTTCTTGGTCTTGGTGCAGTT 1020  
 QY 341 SerAlaIaIaValMetSerSerAlaAspSerSerIIeIeuSerAlaSerMetPheAla 360  
 Db 1021 TCTGCTGCTGATATGATCAGCATCTTCTTCATCTTGTACAGAGATTCAGATTTGCA 1080  
 QY 361 ArgAsnIIeYrGIYIeuSerPheArgGInAsnAlaSerAspIYSGIuIleValaIYrVala 380  
 Db 1081 CGGAAATCATCACAGCTTCTTCATCAAAATGCTTCGAGCAAAAGAAATCGTTTGGGTT 1140  
 QY 381 MetArgIIeThraIeIeuValaPheGIYAlaSerAlaThraIaMetAlaIeuIeuThryI 400  
 Db 1141 ATGCGAATCAACAGTGTGTGTGGAGCATTCGAAACAGCCATGGCTTGTGTCAGAGAA 1200

QY 401 ThrValTYrGIYLeuTYrIleuSerSerAspIleuValTYrIIeIleIlePheProGIn 420  
 Db 1201 ACTGTGTATGGGCTCTGGTACTCACTTCTGACTGTTTACATCTGTTATCTTCCCCCAG 1260  
 QY 421 LeuLeuCysValIeuPheIIeIYSGIYThraThrTYrGIYAlaValaIaGIYrIle 440  
 Db 1261 CTGCTTGTGTACTCTTGTGTAAAGGAAACCAACACTATGGGGCCGTGGCAGGTTATGTT 1320  
 QY 441 PheGIYLeuPheIeuArgIIeThrGIYGIYIuProTYrIeuTYrIeuGInProIle 460  
 Db 1321 TCTGGCTTCTCTGAGATTAACGTGAGGGGAGCCATCTGTATCTTCAAGCTTGATC 1380  
 QY 461 PheTYrProGIYTYrTYrSerAspIYsAsnGIYIleTYrAsnGInArgPheProPheIY 480  
 Db 1381 TTCTACCTGGCTATTTACCTGATGATATGATATATATATATATATATATATATATAT 1440  
 QY 481 ThrIeuSerMetValThraSerPheThraAsnIIeCysValSerTYrIleuAlaIaYr 500  
 Db 1441 AACCTTGCATGATTCATCATTCCTTAACCAACATTTGCATCTCTATCTAGCAAGTAT 1500  
 QY 501 LeuPheGIYSerGIYThrIeuProProIYsIleuAspValaPheAspAlaValaIaArg 520  
 Db 1501 CTATTTGAAAGTGGAACTTGGCACCTTAATTAAGATGTATTTGATGCTGTGTGGCAAG 1560  
 QY 521 HIsSerGIYIuAsnMetAspIYsThrIIeIeuValaIaArgAsnGInIleIYsIleuAsn 540  
 Db 1561 CACAGTGAAGAAACATGTGATTAAGACATTTCTGTCAAAATGAAATTAATTAATTAAT 1620  
 QY 541 GluIleuAlaProValIYsProArgIuSerIeuThraIeuSerSerThraPheThraIYs 560  
 Db 1621 GAACCTTGCACTTGTGAAGCAGACAGACAGACAGACCTCAGCTCACTTACCAATTA 1680  
 QY 561 GluAlaIeIeuLeuAspValaIaSPSerSerProGIYGIYSerGIYThrGIYsPheIeuGIn 580  
 Db 1681 GAGGCTTCTCTGATGTTGATTCAGTCCAGAAAGGCTTGGGACTGAGATTAATTAACAG 1740  
 Db  
 RESULT 12  
 ABX94338  
 ID ABX94338 standard; cDNA, 1743 BP.  
 AC  
 XX  
 DT 13-JUN-2003 (first entry)  
 XX  
 DE Human cDNA encoding high affinity choline transporter, HACT.  
 XX  
 KW Human; ss; gene; HACT; high affinity choline transporter; pain;  
 KW neurotransmitter biosynthesis; learning and memory; aging; epilepsy;  
 KW neurological disorder; spasticity; myoclonus; muscle spasm;  
 KW muscle hyperactivity; stroke; head trauma; neuronal cell death;  
 KW multiple sclerosis; spinal chord injury; dystonia; Alzheimer's disease;  
 KW Myasthenia Gravis; multi-infarct dementia; AIDS dementia;  
 KW Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;  
 KW ALS; attention deficit disorder; organic brain syndrome; schizophrenia;  
 KW nicotine addiction; memory disorder; cognitive disorder.  
 XX  
 OS Homo sapiens.  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 1..1743  
 FT /tag= a  
 FT /product= "HACT"  
 PN  
 XX  
 PD US6500643-B1.  
 XX  
 PF 31-DEC-2002.  
 XX  
 PR 07-SEP-2000; 2000US-00657252.  
 XX  
 PR 07-SEP-2000; 2000US-00657252.  
 PA (UYFL ) UNIV FLORIDA.

XX Wu D, Gu Y, Millard WJ, He Y;  
 XX WPI: 2003-361535/34.  
 DR P-PsDB; AB089879.  
 XX  
 XX Novel isolated polynucleotide (I) that encodes high affinity choline  
 PT transporter protein, useful for preventing, treating or ameliorating  
 PT neurological and cognitive disorders such as Alzheimer's or Parkinson's  
 PT disease.

Claim 2; Col 17-21; 20pp; English.

XX The invention relates to an isolated polynucleotide which encodes a high  
 CC affinity choline transporter (HACT) protein appearing as AB089879. Also  
 CC included are a polynucleotide encoding a fragment consisting of at least  
 CC about 50 amino acids of the HACT protein, a vector comprising the  
 CC polynucleotide, a composition comprising a vector comprising a  
 CC polynucleotide which comprises at least about 12 contiguous nucleic acids  
 CC of a polynucleotide appearing as ABX94339 (encoding choline  
 CC acetyltransferase), a recombinant host cell which comprises the vector  
 CC (used to express the HACT protein or fragment). The polynucleotide is  
 CC useful as a probe or primer to detect the presence of HACT polynucleotide  
 CC in a sample, such as a biological sample, or for screening for test  
 CC agents which bind to the polynucleotide. A pharmaceutical composition  
 CC comprising the polynucleotide is useful for preventing, treating or  
 CC ameliorating neurological and cognitive disorders e.g. pain, spasticity,  
 CC myoclonus, muscle spasm, muscle hyperactivity, epilepsy, stroke, head  
 CC trauma, neuronal cell death, multiple sclerosis, spinal chord injury,  
 CC dystonia, Alzheimer's disease, myasthenia gravis, multi-infarct  
 CC dementia, AIDS dementia, Parkinson's disease, Huntington's disease,  
 CC amyotrophic lateral sclerosis (ALS), attention deficit disorder, nicotine  
 CC addiction, organic brain syndromes, schizophrenia or memory and cognitive  
 CC disorders. HACT is thought to be the rate limiting step in cholinergic  
 CC neurotransmitter biosynthesis and the regeneration (cholinergic transmissions  
 CC are crucial to brain functions such as learning and memory). The present  
 CC sequence encodes human HACT

XX Sequence 1743 BP; 411 A; 395 C; 405 G; 532 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 1,84e-292 Length: 1743  
 Score: 2795.00 Matches: 537  
 Percent Similarity: 96.6% Conservative: 23  
 Best Local Similarity: 92.6% Mismatches: 20  
 Query Match: 93.4% Gaps: 0  
 DB: 9 Indels: 0

US-10-724-806-4 (1-580) x ABX94338 (1-1743)

QY 1 MetProPheHISValGluGlyLeuValAlaIleIleuPheTyrLeuLeuIlePheLeu 20  
 DB 1 ATGGCTTCCATGTCGGAAGAGACTGATACATCAACGCTGCTTCTATATTTTGGCTG 60  
 QY 21 ValGlyIleTTPAlaAlaATripLysThrLysAsnSerGlyAsnProGluGluArgSerGlu 40  
 DB 61 GTTGGAAATATGGCTGCTGCTGGAACCAAAACACATGGGCGCAAGAGCGCGCA 120  
 QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
 DB 121 GCCATCATAGTTGGTGGCGGAGATATTGGTTATTTGGTTGGATTTACATGACACT 180  
 QY 61 ThrTTPValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
 DB 181 ACCGGGCGGAGGAGGGATATATCAATGACACGCTGAAGCAGTTATGATCAAGTTAT 240  
 QY 81 GlyLeuAlaTTPAlaHISAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
 DB 241 GGCTTAGCTTGGGCTCAGGACCAATTTGATATTTCTTAGCTGATTTTAAAGTGCGCTG 300  
 QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120  
 DB 301 TTCTTTGGAAAAACCATATGCTTCAAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCAA 360

QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
 DB 361 ATCTATGAAAAACGATGGCGGACCTCGTTTATTTCTCGACATGATGGGAGAAATGTTTC 420  
 QY 141 ThrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 DB 421 TGGGCTGAGCAATTTCTGCTTGGAGACACACATGAGTGATATGATGATGAT 480  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
 DB 481 ATGCACATTTCTGTCATCATCTCTGCACTCATTTGCCACTCTGTACACACTGGTGAAGG 540  
 QY 181 LeuTyrSerValAlaTyrThrAspValAlaGluLeuPheCysIlePheIleGlyLeuTyr 200  
 DB 541 CTCTATTTCTGGCGCTTACACGATGTCCTTCAAGCTTTTTCATTTTGTAGGGGCTGGG 600  
 QY 201 IleserValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 DB 601 ATCAAGCTCCCTTTGCAATGTCATCTGAGTCGAGTCGAGATCGGCTTCACTGCTGTG 660  
 QY 221 HisAlaLysTyrGlnSerProTrrPheGlyThrIleGlnSerValGluValTyrThrTrr 240  
 DB 661 CATGCCAAATGCCAAAGCCCGTGGGAACTGTGACTCATCTGAAGTCACTACTTGG 720  
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrPheGlnAlaTyrPheGlnArg 260  
 DB 721 CTGTATATTTTTCGTTGTTTATGTCGTGGGTGAATCCATGAGCAAGCATCTTTCAGAGG 780  
 QY 261 ValIleuSerSerSerSerAlaThrTyrAlaGlnValIleuSerPheLeuAlaIlePheGly 280  
 DB 781 GTTCTCTCTTCTTCTCTGACCACTTATGCTCAAGTGTCTTCTGAGCTTTCGGG 840  
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300  
 DB 841 TGCCGTGGATGGCCATCCAGCCATCACTATTGGGGCATTTGGAGCCTCCACAGACTGG 900  
 QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu 320  
 DB 901 AACCAAGCTGATATGGGCTTCCAGATCCCAAGCATACAGAAAGGCGCAGCATGTATTTTA 960  
 QY 321 ProIleValIleuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
 DB 961 CCAATTGTTCTGCAATATCTGCTGCTGTGTAATTTCTTCTTGGCTTGGTGACATT 1020  
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
 DB 1021 TCTGCTGCTGTATGATATGATATGACAGATTTCTTCATCTTGTACAGCAAGTTCCATGTTTGCA 1080  
 QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrrPAl 380  
 DB 1081 CGGAACATCTACAGCTTTCCTTCAGACAAATGCTTCGGCAAGAAATGTTGGGCTT 1140  
 QY 381 MetArgIleThrValIleLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
 DB 1141 ATGGCAATCAAGTGTGTTGTGTTGGACATCTGCAACGCCATGAGCCTTCTGACGAA 1200  
 QY 401 ThrValTyrGlyLeuTrrPyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
 DB 1201 ACTGATGATGGGCTCTGGTACTCAGTTCTGACCTGTTTACATCGTATCTTCCCCAG 1260  
 QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
 DB 1261 CTGCTTTGTGTACCTTTGTTAAGGGAACCAACCATATGGGGCGCTGGCGAGTTATGTT 1320  
 QY 441 PheGlyLeuPheLeuAspGlyIleThrGlyGlyLysProTrrLeuTyrLeuGlnProLeuIle 460  
 DB 1321 TCTGAGCTCTTCTGAGATATCTGAGAGGAGCAATATCTGATCTTCAAGCCCTTGATTC 1380  
 QY 461 PheTyrProGlyTyrTrrPyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
 DB 1381 TTCTTACCTGCTATTTTACCCTGATGATATGATATATATATATGAAATTTCCATTTTAA 1440

QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrIleuAlaIleTyr 500  
DB 1441 AACCTTGCCATGGTTCATCATCTTAAACCAACATTGGATCTCCATCATGACCAAGTAT 1500  
QY 501 LeuPheGluSerGlyThrLeuProPheGluAspValPheAspAlaValAlaIleArg 520  
DB 1501 CTATTGGAAAGTGAACCTTGCACCTTAATGATGTATTGATCTGTTGGAAAG 1560  
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValAlaAsnGluAsnIleuLeuAsn 540  
DB 1561 CACAGTGAAGAAACATGAGTAAAGACATTTCTTGTCAAAATGAAATTTAAATTAGAT 1620  
QY 541 GlnLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
DB 1621 GAACCTTGCACTGTGTGAAGCAGACAGACAGATACCTCAGCTCACTTACCAATTA 1680  
QY 561 GlnAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
DB 1681 GAGGCTTCTCTGATGTGATTCACGTCAGAGAGGCTGGGAGCTGAGATTAATTACAA 1740

RESULT 13  
ADD50638  
ID ADD50638 standard; cDNA, 1743 BP.  
XX  
AC ADD50638;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE cDNA encoding human high-affinity choline transporter (hCHT).  
XX  
KW Human; high-affinity choline transporter; hCHT; chromosome 2q12;  
KW cholinergic function; Parkinson's disease; Huntington's disease;  
KW Alzheimer's disease; schizophrenia; dysautonomia; myasthenia gravis;  
KW brain; cholinergic signalling; antiparkinsonian; anticonvulsant;  
KW neurotropic; neuroprotective; neuroleptic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1743  
FT /tag= a  
FT /product= "hCHT"  
XX  
PN US2003114399-A1.  
XX  
PD 19-JUN-2003.  
XX  
PF 23-JUL-2001; 2001US-00911077.  
XX  
PR 23-JUL-2001; 2001US-00911077.  
XX  
PA (BLAK/) BLAKELY R D.  
PA (APPA/) APPARUNDARAM S.  
PA (FERG/) FERGUSON S.  
XX  
PI Blakely RD, Apparsundaram S, Ferguson S;  
XX  
DR MPI, 2003-810914/76.  
DR P-PSDB; ADD50639.  
XX  
PT Novel isolated polynucleotide encoding human or mouse high affinity  
PT choline transporter polypeptide, useful in gene therapy to increase  
PT cholinergic function in a cell of a patient suffering from Alzheimer's  
PT disease.  
XX  
PS Claim 2; SEQ ID NO 1; 74bp; English.  
XX  
CC The present invention relates to the isolation of polynucleotide  
CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding mCHT, is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic

CC function in the cell that is in a patient having Parkinson's disease,  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signalling. The present sequence encodes hCHT. Note: The sequence data  
CC for this patent was obtained in electronic format directly from the USPRO  
CC web site at [seqdata.uspro.gov](http://seqdata.uspro.gov).  
XX  
SQ Sequence 1743 BP; 412 A; 393 C; 406 G; 532 T; 0 U; 0 Other:  
  
Alignment Scores:  
Pred. No.: 1,84e-292 Length: 1743  
Score: 2795.00 Matches: 537  
Percent Similarity: 96.6% Conservative: 23  
Best Local Similarity: 92.6% Mismatches: 20  
Query Match: 93.4% Indels: 0  
DB: Gaps: 0  
  
US-10-724-806-4 (1-580) x ADD50638 (1-1743)  
  
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrIleuLeuIlePheLeu 20  
DB 1 ATGGCTTTCATGTGGAGAGACTGATAGCTATCATGTGTTCACTTCAATTTGCTG 60  
QY 21 ValGlyIleTyrAlaIleTyrLysThrLysAsnSerGlyAsnProGluGluAspSerGlu 40  
DB 61 GTTGGAAATAGGGCTGCTGGAGAACCCAAACAGTGGACGCCAGAGCCGACGGA 120  
QY 41 AlaIleIleValAlaGlyIleArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAla 60  
DB 121 GCATCATATGTTGTGTGGCCGAGATATGTTATGTTGTGTGATTTACATGACAGCT 180  
QY 61 ThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
DB 181 ACCTGGGTGGAGAGAGGTATATCAATGACAGCTGAAGCACTTATATGACAGGTTAT 240  
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
DB 241 GGCTTAGCTTGGGCTAGGACCAATGGATATCTCTTAGTGAATTTAGGTGGCTG 300  
QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120  
DB 301 TTCTTTGCAAAACCTATGCGTTCAAGGGGTATGTCACCATGTATGACCGTTTCAAGCA 360  
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyIleMetPhe 140  
DB 361 ATCTATGAGAAACGCAATGGCGGACTCTGTTATCTCGACTGATGGAGAAATGTTTC 420  
QY 141 TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
DB 421 TGGGCTGACAGAAATTTCTCTGCTTGGAGACCATGACGCTGATCATCGATGGAT 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
DB 481 ATGCACATTTCTGTATCATCTCTGTCACATTCATGCACTGTACACACGTGGAGGG 540  
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
DB 541 CTCTATTCTGTGGCCACATGATGTCGTCACCTTTTGGCAATTTTGTAGGGCTGTG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
DB 601 ATCAGGCTCCCTTTCATGATTCACATCTGACAGTGCAGACATCGGTTCACTGCTGTG 660  
QY 221 HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGluSerValGluValTyrThrTyr 240  
DB 661 CATGCCAAATATACCAAAAGCCGTGGGAGACGTGATCTGAGCTCATCTACTCTTGG 720  
QY 241 LeuAspAsnPheLeuLeuMetLeuGlyIleProTyrGlnAlaTyrPheGlnArg 260  
DB 721 CTTGATATGTTTCTGTTGTGATGCTGGGTGAGATCCATGCGACATATCTTCAAGG 780



QY 41 AAlaIleIValGIyGIyArGAspIleGIyLeuIleuValGIyGIyPheThrMetThrAla 60  
 DB 121 GCCATCATGTTGGTGGCCGAGATATTGGTTATTGGTTGGTGGATTTACCATGACGACT 180  
 QY 61 ThrTPValIGIyGIyGIyTyrIleAsnGIyThrAlaGIuAlaValIyrgIyProGIyCys 80  
 DB 181 ACCGTGGGTGGAGAGGTATATCAATGACACAGCTGAAGCAGTTTATATGACCGCTTAT 240  
 QY 81 GIyLeuAlaTPValaHisAlaProIleGIyTyrSerIleuSerIleuGIyGIyLeu 100  
 DB 241 GGGCTTACCTGGGCTAGGACCAATGGATTTCTTATGTTGATTTAGGTGGCTGG 300  
 QY 101 PhePheAlaIySPromerArGSerIySGIyTyrValIThrMetLeuAspPropheIySGIn 120  
 DB 301 TTCTTTGCAAAACCTATGGCTTCAAGGGGTATGACCATGTTAACCCTGTTTACGCA 360  
 QY 121 IIEYrGIyLyahrgMetGIyGIyLeuLeuPheIleProAlaLeuMetGIyIuMetPhe 140  
 DB 361 ATCTATGGAAAAAGCATGGGCGGACCTCTGTTTATCTGCACTGATGGGAGAAATGTTT 420  
 QY 141 TPValaAlaIlePheSerAlaLeuGIyAlaThrIleSerValIleIleAspValaSP 160  
 DB 421 TGGGCTGACGACATTTCTCTGCTTGGAGACCATGACGCTGATCATCGATGGAT 480  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleuTyrThrIleuValGIyGIy 180  
 DB 481 ATGCACATTTCTGTCTCATCTCTGCACTCATTCGCACTGTGACACACGCTGGGAGAG 540  
 QY 181 LeuTyrSerValAlaIyThrAspValValGIuLeuPheCysIlePheIleGIyLeuTP 200  
 DB 541 CTCTATTTCTGTGGCTTACACTGATGTCTGCTTGGCTTTTGGCAATTTTGGGCGCTGG 600  
 QY 201 IIEserValProPheAlaIleuSerHisProAlaValIThrAspIleGIyPheThrAlaVal 220  
 DB 601 ATCAGGCTCCCTTTCATATGTCATCCGACGACGACATCGAGTCTGCTGCTG 660  
 QY 221 HisAlaIySerGIyInserProTPLeuGIyThrIleGIySerValGIuAlaIyThrTP 240  
 DB 661 CATGCCAATATACAAAGCGGTGGCTGGGAGCTGTGACTCATCTGAAGCTTCTTGG 720  
 QY 241 LeuAspAsnPheLeuLeuLeuMetIleuGIyIleProTPGIuAlaIyThrPheGIuArg 260  
 DB 721 CTTGATAGTTTCTGTGTGATGCTGGGTGAATCCCATGCAACATATCTTTCAGAG 780  
 QY 261 ValIleuSerSerSerAlaThrIyrrAlaGIuValIleuSerPheIleuAlaIlePheGIy 280  
 DB 781 GTTCTCTCTTCTTCCACGACCATATGCTCAAGTCTGTCTTCCCTGGAGCTTTCGG 840  
 QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGIyAlaIleGIyAlaIleSerThrAspTP 300  
 DB 841 TGGCTGTATGGCCATCCAGCCATATCATTCATTTGGGGCATTTGGAGCATCAAGACTGG 900  
 QY 301 AsnGIuThrAlaIyGIyTyrProAspProIySThrIySGIuGIuAlaAspMetIleu 320  
 DB 901 AACCCAGACTGCATATGGCTTCCAGATCCCAAGACTACAGAAAGGACGACATGATTTTA 960  
 QY 321 ProIleValIleuGIuThrIleuCysProValIyrrIleSerPhePheGIyLeuGIyAlaVal 340  
 DB 961 CCAATGTCTGAGATCTCTGCTGCTGTATATTTCTTCTTGGTGTGAGAT 1020  
 QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleuSerAlaSerSerMetPheAla 360  
 DB 1021 TCTGCTGCTGTTATATGATCATGACAGATTTCTTCATCTTGTGACAGCATTTGCA 1080  
 QY 361 ArgAsnIleTyrGIuLeuSerPheArgGIuAsnAlaSerAspIySGIuIleValTPVal 380  
 DB 1081 CGGAACATCTACACAGCTTCTTCCAGCAAAATGCTTCGCAAAATAATCGTTTGGGTT 1140  
 QY 381 MetArgIleThrValIleuValPheGIyAlaSerAlaThrAlaMetAlaLeuLeuThrIyS 400  
 DB 1141 ATGCGAATCAAGTGTGTTGGTGTGGAGCATCTGCACAGCCATGGCTTGTGACGAAA 1200  
 QY 401 ThrValIyrgIyLeuTPyrrIleuSerSerAspLeuValIyrrIleIlePheProGIu 420

DB 1201 ACTGTATAGGCTCTGTACTCAAGTCTGCACCTGTATTACATCGTTATCTTCCCCAG 1260  
 QY 421 LeuLeuCysValLeuPheIleIySGIyThrAsnThrTyrGIyAlaAlaIaGIyTyrIle 440  
 DB 1261 CTGCTTGTGTACTCTTGTGTAAAGGAGCAACACTATAGGGCCGTGGCAGGTTATGTT 1320  
 QY 441 PheGIyLeuPheLeuArgIleThrGIyGIyIuProTyrLeuTyrIleuGIuProIle 460  
 DB 1321 TCTGGCTTCTTCTGAGATTAACCTGAGGGAGCCATATCTGTATCTTACGCCCTTGATC 1380  
 QY 461 PheTyrProGIyTyrTyrSerAspIySAsnGIyIleTyrAsnGIuArgPhePropheIyS 480  
 DB 1381 TTCTACCTGCTATTAACCTGATGATATGATATATATATATATATATATATATATAT 1440  
 QY 481 ThrLeuSerMetValIThrSerPhePheThrAsnIleCysValIleSerTyrIleuAlaIyS 500  
 DB 1441 ACACTTGCATGTGTATCATCATTTCTTAAACCAACATTTGATCTCTATCTAGCAAGTAT 1500  
 QY 501 LeuPheGIuSerGIyThrIleuProProlYsLeuAspValPheAspAlaValAlaIArg 520  
 DB 1501 CTATTTGAAAGTGAACCTTGGCACTTAATATATATATATATATATATATATATAT 1560  
 QY 521 HisSerGIuGIuAsnMetAspIySThrIleLeuValArgAsnGIuAsnIleIySLeuAsn 540  
 DB 1561 CACAGTGAAGAAACATGATTAAGACAAATCTTGTCTCAAAAATGAAATATTAATTAAT 1620  
 QY 541 GIuLeuAlaProValIySPromerArgInserIleuThrIleuSerSerThrPheThrAsnIyS 560  
 DB 1621 GAATTCGACCTTGTGAAGCAGACAGACGATGACCTTCAAGCTCAACTTTCACCAATTA 1680  
 QY 561 GIuAlaLeuLeuAspValaIleAspSerProGIuGIySerGIyThrGIuAspAsnLeuGIu 580  
 DB 1681 GAGGCTTCTTGTATGTATTTCCAGTCCAGAAAGGCTTGGGACTGAAGATATTTTACAG 1740  
 DB  
 RESULT 15  
 ADD50646  
 ID ADD50646 standard; DNA; 1813 BP.  
 AC  
 ADD50646;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE High-affinity choline transporter (CHT) associated DNA sequence #2.  
 XX  
 KW High-affinity choline transporter; CHT; cholinergic function;  
 KW Parkinson's disease; Huntington's disease; Alzheimer's disease;  
 KW schizophrenia; dysautonomia; myasthenia gravis; brain;  
 KW cholinergic signalling; antiparkinsonian; anticonvulsant; nootropic;  
 KW neuroprotective; neuroleptic; ds.  
 OS  
 OS Unidentified.  
 PN US2003114399-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 23-JUL-2001; 2001US-00911077.  
 XX  
 PR 23-JUL-2001; 2001US-00911077.  
 XX  
 PA (BLAK/) BLAKELY R D.  
 PA (APPA/) APPARISUNDARAM S.  
 PA (FERG/) FERGUSON S.  
 PI BlakeLY RD, Apparsundaram S, Ferguson S;  
 DR MPI; 2003-810914/76.  
 XX  
 PT Novel isolated polynucleotide encoding human or mouse high affinity  
 PT choline transporter polypeptide, useful in gene therapy to increase  
 PT cholinergic function in a cell of a patient suffering from Alzheimer's  
 PT disease.

XX Disclosure, SEQ ID NO 9; 74bp; English.

XX The present invention relates to the isolation of polynucleotide  
CC sequences encoding human and mouse high-affinity choline transporter  
CC (hCHT and mCHT respectively), and the proteins they encode. The gene  
CC encoding hCHT is located on chromosome 2q12. The polynucleotide sequence  
CC encoding hCHT is useful for expressing hCHT recombinantly. The hCHT  
CC polynucleotide sequence when delivered to a cell, increases cholinergic  
CC function in the cell that is in a patient having Parkinson's disease,  
CC Huntington's disease, Alzheimer's disease, schizophrenia, dysautonomia or  
CC myasthenia gravis. The hCHT antibody is useful for controlling  
CC transporter CHT proteins to the brain, and for treating the above  
CC mentioned diseases. The antibody is also useful for diagnosing the above  
CC mentioned disorders and to detect the influence of cholinergic  
CC signaling. The present DNA sequence of unknown function is provided in  
CC the electronic sequence data but is not mentioned in the printed  
CC specification. Note: The sequence data for this patent was obtained in  
CC electronic format directly from the USPTO web site at [seqdata.uspto.gov](http://seqdata.uspto.gov).

XX Sequence 1813 BP; 440 A; 406 C; 417 G; 550 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1,966-2392	Length:	1813
Score:	2795.00	Matches:	537
Percent Similarity:	96.6%	Conservative:	23
Best Local Similarity:	92.6%	Mismatches:	20
Query Match:	93.4%	Indels:	0
DB:	10	Gaps:	0

US-10-724-806-4 (1-580) x ADD50646 (1-1813)

QY 1 MetPropheni:sva:lg:ly:leu:val:ala:leu:leu:phe:tyr:leu:leu:leu:leu: 20  
DB |ATGGCTTCCAGTGGAGAGACTGATGCTATCATGCTTCTTACTTCTTAATTTTCTG 78  
QY 21 ValGlyIleTrrAlaAlaTrrLysThrLysAsnSerGlyAsnProGluGlnArgSerGlu 40  
DB ||||||TrrAlaAlaTrrLysThrLysAsnSerGlyAsnProGluGlnArgSerGlu 40  
DB 79 GTTGAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 138  
QY 41 AlaIleIleValAlaGlyAlaArgAspIleGlyLeuLeuValAlaGlyPheMetThrAla 60  
DB |GCCATCATAGTTGGTGGCGGAGATATGTTTATGTTGGTGGATTTACATGACAGT 198  
QY 61 ThrTrrValAlaGlyAlaGlyTrrIleAsnGlyThrAlaGlnAlaValAlaGlyProGlyCys 80  
DB |ACCTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 258  
QY 81 GlyLeuAlaTrrAlaAlaIleAlaProIleGlyTrrSerLeuSerLeuIleGlnGlyLeu 100  
DB ||||||TrrAlaAlaIleAlaProIleGlyTrrSerLeuSerLeuIleGlnGlyLeu 100  
QY 259 GAGCTAGCTTGGGCTCAGGACCAATGATGATCTTATGCTTAATTTTATGAGGAGGCTG 318  
QY 101 PhePheAlaLysProMetArgSerLysGlyTrrValAlaThrMetLeuAspProPheLysGln 120  
DB |TTCCTTGGAAACCTATGCTGCTCAAGGGTATGAGACATGTTAGACCCGTTTACAGCA 378  
QY 319 TTTCTTGGAAACCTATGCTGCTCAAGGGTATGAGACATGTTAGACCCGTTTACAGCA 378  
QY 121 IleTrrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140  
DB |ATCTATGAGAAACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 438  
QY 141 TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaTrrIleSerValIleIleLysValAsp 160  
DB |TGGGCTGAGCAATTTCTCTGCTTGGAGGACCATCAGGATATCATCATGATGAT 498  
QY 439 TGGGCTGAGCAATTTCTCTGCTTGGAGGACCATCAGGATATCATCATGATGAT 498  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTrrThrLeuValAlaGly 180  
DB |ATGCACATTTCTGTCATCATCTCTGCACATCTTGCACATCTGTACACACTGGTGGAGGG 558  
QY 499 ATGCACATTTCTGTCATCATCTCTGCACATCTTGCACATCTGTACACACTGGTGGAGGG 558  
QY 181 LeuTrrSerValAlaAlaTrrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrr 200  
DB |CTCTATTTCTGGGCTTACATGATGCTGTTGAGCTCTTTTGCATTTTGTAGGGGCTGG 618  
QY 559 CTCTATTTCTGGGCTTACATGATGCTGTTGAGCTCTTTTGCATTTTGTAGGGGCTGG 618  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValAlaThrAspIleGlyPheThrAlaVal 220

DB 619 ATCAGCGCTCCCTTGGATTTGTCATCTGAGCTGCGAGCATTCGGGTTTACATGCTGTG 678  
QY HisAlaLysTrrGlnSerProTrrPheGlyTrrIleGlnSerValAlaValAlaTrrThrTrr 240  
DB |CATGCGAAATATCAAAAGCCGTGGCTGGAGACTGTGATCTGATCTGAAGTCTACTCTTGG 738  
QY 241 LeuAspAsnPheLeuLeuLeuMetIleGlyIleProTrrGlnAlaTrrPheGlnArg 260  
DB |CTGTATATTTTCTGTTGTTATGCTGGGTGGAATCCATGAGCAAGCTTCTTACAGAGG 798  
QY 261 ValLeuSerSerSerSerAlaThrTrrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280  
DB |GTTCTCTTCTTCTTCTTCTGACCACTATGCTCAAGTGTCTTCTTCTTCTTCTTCTTCTG 858  
QY 799 GTTCTCTTCTTCTTCTTCTGACCACTATGCTCAAGTGTCTTCTTCTTCTTCTTCTTCTG 858  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300  
DB |TGCCTGGATGGATGCCATCCAGCATCTCATTTGGGCGCATTTGAGCATCAACAGACTGG 918  
QY 859 TGCCTGGATGGATGCCATCCAGCATCTCATTTGGGCGCATTTGAGCATCAACAGACTGG 918  
QY 301 AsnGlnTrrAlaTrrGlyTrrProAspProLysThrLysGlnGlnAlaAspMetIleLeu 320  
DB |AACCAGACTGATATGAGGCTTCCAGATCCCAAGCTACAGAAAGAGGACAGATGATTTTA 978  
QY 321 ProIleValLeuGlnTrrLeuCysProValTrrIleSerPhePheGlyValAlaVal 340  
DB |979 CCAATGTTTGGAGATATCTGCTGCTGTGATATTTCTTCTTCTTCTTCTTCTTCTTCTG 1038  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
DB |TCTGCTGCTGTTATGTCATTCAGCAGATCTTCCATCTTGTGACAGATGTTCCATGTTTCCA 1098  
QY 361 ArgAsnIleTrrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrrVal 380  
DB |CGGAACATCTACACAGCTTCTTCCAGCAAAATGCTTGGACAAAGAAATGTTGGGCTT 1158  
QY 1099 CGGAACATCTACACAGCTTCTTCCAGCAAAATGCTTGGACAAAGAAATGTTGGGCTT 1158  
QY 381 MetArgIleTrrValLeuValPheGlyAlaSerAlaTrrAlaMetAlaLeuLeuThrLys 400  
DB |ATGCAATCAACAGTGTGTTGTTGTTGAGCATCTGCACAGCATGAGCTTGTCTGACGAAA 1218  
QY 1159 ATGCAATCAACAGTGTGTTGTTGTTGAGCATCTGCACAGCATGAGCTTGTCTGACGAAA 1218  
QY 401 ThrValTrrGlyLeuTrrTrrLysSerSerAspLeuValTrrIleIleIlePheProGln 420  
DB |ACTGATGATGGGCTGTGGTACTGATCTGATCTGACCTGTTTATCATGTTATCTTCCCCAG 1278  
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTrrGlyAlaValAlaGlyTrrIle 440  
DB |CTGCTTGTGATCTTGTGTTGAGGAGCAACACTTGGGCGGCTGAGGTTATGTT 1338  
QY 1279 CTGCTTGTGATCTTGTGTTGAGGAGCAACACTTGGGCGGCTGAGGTTATGTT 1338  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTrrLeuTrrLeuGlnProLeuIle 460  
DB |TCTGGCTCTTCTCTGAAATATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1398  
QY 1339 TCTGGCTCTTCTCTGAAATATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1398  
QY 461 PheTrrProGlyTrrTrrSerAspLysAsnGlyIleTrrAsnGlnArgPheProPheLys 480  
DB |TTCATACCTGGCTATTTACCTGATGATATGATATATTAATTCAGAAATTTCCATTTAA 1458  
QY 1399 TTCATACCTGGCTATTTACCTGATGATATGATATATTAATTCAGAAATTTCCATTTAA 1458  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTrrLeuAlaLysTrr 500  
DB |ACACTTGGCAATGTTATCATCTTCTTAACCAACTTTCACCTCTCATCTGACCAAGAT 1518  
QY 1459 ACACTTGGCAATGTTATCATCTTCTTAACCAACTTTCACCTCTCATCTGACCAAGAT 1518  
QY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaLysArg 520  
DB |CTATTTGAAAGTGAACCTTCCACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1578  
QY 1519 CTATTTGAAAGTGAACCTTCCACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1578  
QY 521 HisSerGlnGlnAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsn 540  
DB |CAGAGTGAAGAAACATGATGATTAAGCAATTTCTTGTCAAAATATTAATTAATTAATTA 1638  
QY 1579 CAGAGTGAAGAAACATGATGATTAAGCAATTTCTTGTCAAAATATTAATTAATTAATTA 1638  
QY 541 GlnLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrLysLys 560  
DB |GAACTTGCACCTTGGAGAACCGACAGAGAGCATGACCTTCACTCAACTTCAACCAATAAA 1698  
QY 1639 GAACTTGCACCTTGGAGAACCGACAGAGAGCATGACCTTCACTCAACTTCAACCAATAAA 1698  
QY 561 GlnAlaLeuLeuAspValAspSerSerProGlnGlySerGlyThrGlnAspAsnLeuGln 580

Tue Jul 11 13:42:51 2006

us-10-724-806-4.p2n.rng

Page 21

Db 1699 GAGCCTTCTGATGTGATTCACGTCAGAAAGGCTGGAAGTGAAGATTAATTACAG 1758

Search completed: July 10, 2006, 15:06:15  
Job time : 910 secs

---

**This Page Blank (uspto)**



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocooperation Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: July 10, 2006, 14:03:36 ; Search time 7655 Seconds  
(without alignments)  
6355.300 Million cell updates/sec

Title: US-10-724-806-4

Perfect score: 2993  
Sequence: 1 MPFHEGVAILFLYLFL.....EALIDVDSPEGGSTEDNLQ 580

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.model -DEV=xlh  
-O=/abs/ABSWEB/spool/US10724806/runat\_10072006\_140248\_7740/app\_query.fasta\_1  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs802h  
-USER=US10724806\_@CN1\_1\_6323\_@runat\_10072006\_140248\_7740 -NCPU=6 -ICPU=3  
-NO MAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_hic:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_gse1:\*  
12: gb\_gse2:\*  
13: gb\_gse3:\*  
14: gb\_gse4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2973	99.3	4097	6 AK053063
2	2973	99.3	4306	6 AK034415 Mus muscu
3	2410	80.5	1743	14 AY413300 Mus muscu
4	2257	75.4	1743	14 AY413298 Homo sapi

5	2251	75.2	1743	14 AY413299	AY413299 Pan trogl
6	1457	48.7	4122	6 AK164116	AK164116 Mus muscu
7	1243	41.5	882	5 CK770440	CK770440 958613 MA
8	1065	35.6	707	5 CD350164	CD350164 UI-M-FYO-
9	1004	33.5	828	14 CT403380	CT403380 Sus scrof
10	988.5	33.0	669	10 DM65738	DM65738 CNB333-A0
11	987	33.0	669	4 BY727598	BY727598 BY727598
12	968	32.3	843	10 DT207246	DT207246 JGI CAAS1
13	964	32.2	781	10 DT197042	DT197042 JGI CAAS5
14	953	31.8	1114	10 DT972151	DT972151 CLJ138-H0
15	890	29.7	1054	10 DM614674	DM614674 CLJ282-B0
16	842.5	28.1	672	14 AG157499	AG157499 Pan trogl
17	828	27.7	834	9 CX840109	CX840109 JGI CAAR9
18	824	27.5	516	7 BE233479	BE233479 139685 MA
19	822.5	27.5	800	1 AL669749	AL669749 AL669749
20	790	26.4	548	9 DA361315	DA361315 DA361315
21	784	26.2	576	9 DN991044	DN991044 TC112868
22	746	24.9	753	10 DV838598	DV838598 LB01121.C
23	746	24.9	801	3 BM002036	BM002036 BM002036
24	719	24.0	1021	10 DM616805	DM616805 CLJ294-F1
25	707.5	23.6	1037	1 AL666817	AL666817 AL666817
26	704	23.5	934	8 CV123740	CV123740 OSTF30017
27	703	23.5	650	7 BB626260	BB626260 BB626260
28	684	22.9	694	8 CX033587	CX033587 134598 N
29	681	22.8	874	5 CK449221	CK449221 892772 MA
30	680.5	22.7	658	3 BM629925	BM629925 170006875
31	676.5	22.6	549	3 BM274870	BM274870 BM274870
32	660.5	22.1	641	2 BI630566	BI630566 RH59836.5
33	659	22.0	682	9 DN604180	DN604180 ACAC-aab8
34	659	22.0	842	5 CK184985	CK184985 EST774300
35	652.5	21.8	640	2 BI629504	BI629504 RH58381.5
36	638.5	21.3	941	5 CD360297	CD360297 AGENCOURT
37	631.5	21.1	813	10 DM569404	DM569404 EST_ssa1
38	630	21.0	675	4 BY729567	BY729567 BY729567
39	629	21.0	658	8 CX196909	CX196909 Sa_mxo_52
40	628	21.0	541	7 AW668962	AW668962 111664 MA
41	624	20.8	808	5 CK184984	CK184984 EST774299
42	619	20.7	576	4 CB391304	CB391304 OSTF149A8
43	619	20.7	624	2 BJ122485	BJ122485 BJ122485
44	600.5	20.1	565	2 BJ125564	BJ125564 BJ125564
45	600.5	20.1	583	3 BM277281	BM277281 BM277281

## ALIGNMENTS

RESULT 1  
AK053063  
LOCUS  
DEFINITION  
Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:D930038E20 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.

ACCESSION  
AK053063.1 GI:26343192  
VERSION  
AK053063.1  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus

REFERENCE  
1 Carninci, P., Hayashizaki, Y., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kire, A. and Hayashizaki, Y.  
RIKEN Integrated sequence analysis (RISA) system--384-Format  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
PUBMED  
11076861

TITLE  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

REFERENCE  
AUTHORS

TITLE  
The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

TITLE  
RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.  
Antisense transcription in the mammalian transcriptome  
Science 309, 1564-1566 (2005)

JOURNAL  
REFERENCE  
AUTHORS

TITLE  
The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).  
The Transcriptional landscape of the mammalian genome  
Science 309, 1559-1563 (2005)

JOURNAL  
REFERENCE  
AUTHORS

TITLE  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa, 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE  
JOURNAL

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/  
Location/Qualifiers

FEATURES  
SOURCE

1. 4097  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:D930038E20"  
/db\_xref="taxon:10090"  
/clone="D930038E20"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="15 days embryo"  
512..2254  
/note="unnamed protein product; putative

CDS

solute carrier family 5 (choline transporter), member 7  
[MD]MG1:1927126 GB|NW\_022025, evidence: BLASTN, 99%,  
match=1743")  
/codon\_start=1  
/protein\_id="BAC35253.1"  
/db\_xref="gi:26343193"  
/translation="MSRHVEGLVALILFLYLILFLVGIAMKTKNSGNPERSEALIV  
GGRIIGLVGFMTATWVGGLYINGDEAVYGGCGGLMAWAQAFGLSLISLILGGLF  
AKPMRSKGYVTMDPFOQIYGRMGGLFLPALGMEMMAAFSALGATISVILIDV  
VNISIVSALIAILYTLVGLYSAVYDVDFICIFIGLMTISVPALSHPAVTDIGFT  
AVHAKYOSPMIGTIESVEVYTWLDFLMLMGGLPMOAYFORVSSSAYAOVLSFL  
AAPGLVMAIPALICIGALIGASTDMONRATYGPDPKTEADAMILPIYQYCPYISF  
FGLGAVSAWWSADSSILSSASMPNARIYGLSTRONASDAEIVYWMVITLVNGASA  
TAMALIKTVYGLWLSLDVYIILFPLGLVLTIGNTYGAAGYIFGLPLRTTGG  
EPYLYQPLIFPGYSPDKNGIYNORPEFKLSWTSFTNIVSVYAKLYFESETLP  
PKLDFPDAVARHSEENMDKTLIVRENENIKINELAPVAPROSLTSLSTFWKEALDV  
DSPEGSETEENLQ"

ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 4097  
Score: 2973.00 Matches: 577  
Percent Similarity: 99.7% Conservative: 1  
Best Local Similarity: 99.5% Mismatches: 2  
Query Match: 99.3% Indels: 0  
DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x AK053063 (1-4097)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheLeuLeuIlePheLeu 20  
DB 512 ATGCTTTCACGTAAGAGACGTGTACCTTATCTCTTCTTCTTCTTCTTCTG 571

QY 21 ValGlyIleTrrPalaIalaTrrPlyThrIysAsnSerGlyAsnProGluIAsnSerGlu 40  
DB 572 GTTGAATATATGGCGTCATGGAACCAAAAACAGCGCAACCCAGAGAGCGAGTGA 631

QY 41 AlaIleIleValGlyGlyIAsnAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
DB 632 GCCATCATATAGCGGGGGCGGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 691

QY 61 ThrTrrPalaIAsnGlyIleGlyIleAsnGlyThrIAsnGlyIleValIleGlyProGlyCys 80  
DB 692 ACCTGGGTGGAGGCGCTCATCATATGAGACAGAGAACGATATAGGCCAGGTGT 751

QY 81 GlyLeuAlaTrrPalaIAsnAlaProIleGlyIleSerLeuSerLeuIleLeuGlyIleu 100  
DB 752 GGTCTAGCTTGGCGTCAGGACCCATGGAATATCTCGAGCTATTTAGGTGTCTG 811

QY 101 PhePheAlaIAsnProMetArgSerIleGlyIleValIleMetLeuAspProPheIleGln 120  
DB 812 TTTTGTGGAAACCTATCGTTCAGGAGATATGACTATGTATGACCCATTTCAACG 871

QY 121 IleTrrGlyIAsnMetGlyIleLeuPheIleProAlaLeuMetGlyIleuMetPhe 140  
DB 872 ATCTATGAAAGCGATGAGGCGGTCTTCTTCACTCGCTGACTATGAGAGATGTTT 931

QY 141 TrrPalaIalaIlePheSerAlaLeuGlyIleAlaThrIleSerValIleIleAspValAsp 160  
DB 932 TGGGTGCGAGCAATTTCTCGCATTTAGGGGCGCACATCAGGAGTATCATGATGAT 991

QY 161 ValAsnIleSerValIleValIleSerAlaLeuIleAlaIleLeuTrrLeuValGlyGly 180  
DB 992 GTGAACATATCGCTATGCTCGCACTCATGCAATTTTAAACCTATGAGGTGG 1051

QY 181 LeuTrrSerValAlaIleTrrAspValValGlnLeuPheCysIlePheIleGlyLeuTrrP 200  
DB 1052 CTCTACTCTGAGCATATATGATGTTGTCCAGCTATTTCTGATTTTATAGACTGTGG 1111

QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220  
DB 1112 ATCAGTGTCCCTTTTGGCCCTGTGACATCTGTGAGTACCGACATCGGATTCACAGCTGTG 1171

QY 221 Hsajaalytyglnserprotpleuglythrilegluservalgluvaltyrthrtp 240  
 DB 1172 CATGTAATAATACAGAGTCCCTGGCTGGGAACCATGAAATCACTGGAAGCTACACTGG 1231  
 QY 241 leuapasnphleuleuleuleuwtleuglygllyleprotptglualaryrphleglnarg 260  
 DB 1232 CTTGAATAATTTCTGTATATGAAGCTGGGGGAGATCCCATGGCAACCTTACCTCCAGAG 1291  
 QY 261 VallenseSerSerSerSerAlaThrYrAlaGlnValIleuserPheleuAlaIlePhegly 280  
 DB 1292 GTTCCTCTTCTCATCTCCAGCCACTTATGCTCAGGTCCTGCTTCTTCTGGCAGCTTTGGG 1351  
 QY 281 Cysleuvalmetalauleuproalailecysilleglyalalleglyalaserthrpastrp 300  
 DB 1352 TGCCTGTGATGGCTCTACCCCGCATATGATGAGAGCTATTGGAGCTTCCACAGACTGG 1411  
 QY 301 AenglnthAlaTyrglyTyrrProappProlysthlysglgluAlaaspMetIleleu 320  
 DB 1412 AACCAAGACGCTACCGGATCCAGATCCCAAGACTAAGAGAGACAGACATGATTTCTC 1471  
 QY 321 ProilevalleuglnTyrrleuCyrrProvalTyrrIleSerPhePheglyleuglyAlaVal 340  
 DB 1472 CCGATGCTTTCGAGACCTCTGCTGTATCATCTCTTCTTGGGCTTGGTCTGTT 1531  
 QY 341 SerAlaAlaValMetSerSerAlaaspSerSerIleleuserAlaserMetPheAla 360  
 DB 1532 TAGCTGCTGTATGCTCCTCAGCTGACTGCTGCTCATCTGTCGCGAGTTCTATGTTGCT 1591  
 QY 361 ArgAsnIleTyrglnleuserPheleuglnAsnAlaserAspIleleuAlaVal 380  
 DB 1592 CGGAATATTCACGACTTCTTCCCTTCAACAAAGATGATGAGCAAGAAATGTGTGGCTC 1651  
 QY 381 MetArgIleThrValleuValPheglyAlaserAlaMetAlaMetAlaMetThrlys 400  
 DB 1652 ATGAGGATACCTGCTTGTGTGTCGAGACATCTGCACACCCATGGCTTCTGCGAGAG 1711  
 QY 401 ThrValTyrglyleuTyrrleuserSerAspIleuValTyrrIleIlelePheProgin 420  
 DB 1712 ACTGTATATGGGCTGTGTAACCTGAGCTGAGCTGCTGTATCATCATCTTCCACAG 1771  
 QY 421 leuleuCyvalleuPheIlelysglyThrAsnThrTyrglyAlaValAlaVal 440  
 DB 1772 CTGCTCTGTACTCTTCAACAAAGAACCAACATTAAGGGCAGTGTGTTATAT 1831  
 QY 441 PheglyleuPheleuArgIleThrIleTyrglygluProTyrrleuTyrrleuIle 460  
 DB 1832 TTGGACATATCTCGAAATTAAGGAGAGAGACCATATCTATATTCAGCCCTTAATC 1891  
 QY 461 PheTyrrProglyTyrrTyrrSerAspIleAsnIleTyrrAsnIleArgPheProPheIys 480  
 DB 1892 TTCTACCGGTTATTAATCTGACCAAGAAATGATATACATCAGAGGTTCCCATTTAA 1951  
 QY 481 ThrIleuserMetValThrSerPhePheThrAsnIleCysValaserTyrrleuAlaIleTyrr 500  
 DB 1952 ACTCTCTCATGTATCCCTCATCTTACCAACATTTGCTTCTTATATCAGCAAGTAT 2011  
 QY 501 leuPheglSerglyThrleuProProlystleuAspValPheAspAlaValAlaArg 520  
 DB 2012 CTAATTTGAAGTGAACCTTGGCTCCCAAAATTAAGATTTATGCTGTGTGGCAAG 2071  
 QY 521 HisSerIleuglnAsnMetAspIleThrIleleuValArgAsnIleAsnIleIleleuAsn 540  
 DB 2072 CACAGGAGAGAGAACATGAGCAAGACCATCTAGTCAGAAATGAATAATCAATTAAT 2131  
 QY 541 GluleuAlaProValysProArgIleuserIleuSerIleuSerIleuPheThrAsnIle 560  
 DB 2132 GAACTTGCACCTGTGAACCTCGGCGAGGCTTACCTCAGTTCACTTCCACCAATTAAG 2191  
 QY 561 GluAlaIleuAspValaspSerSerProgluglySerIleTyrrgluAspAsnleuGln 580  
 DB 2192 GAGGCCCTCTTGTATGTTGATTTCCATCCGAGAGGGGTCTGGGACTGAAGATACTTACAA 2251

AK034415  
 LOCUS  
 DEFINITION  
 AK034415 4306 bp mRNA linear HTC 02-SEP-2005  
 Mus musculus adult male diencephalon cDNA, RIKEN full-length  
 enriched library, clone:933018BK24 product:solute carrier family 5  
 (choline transporter), member 7, full insert sequence.  
 AK034415  
 VERSION  
 AK034415.1 GI:26329926  
 KEYWORDS  
 HTC; CAP trapper.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.  
 1  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 2  
 10349636  
 3  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 11042159  
 4  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Taahito, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 11076861  
 4  
 THE RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)  
 5  
 THE FANTOM Consortium, the RIKEN Genome Exploration Research Group  
 Phase I and II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6  
 RIKEN Genome Exploration Research Group, Genome Science Group  
 (Genome Network Core Team) and the FANTOM Consortium.  
 Antisense transcription in the Mammalian Transcriptome  
 Science 309, 1564-1566 (2005)  
 7  
 THE FANTOM Consortium, Riken Genome Exploration Research Group and  
 Genome Science Group (Genome Network Project Core Group).  
 The transcriptional landscape of the Mammalian Genome  
 Science 309, 1559-1563 (2005)  
 8  
 (bases 1 to 4306)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohse, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of

RESULT 2

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedic Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>,  
URL: <http://fantom.gsc.riken.jp/>,  
Location/Qualifiers

## FEATURES

## Source

1..4306  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:9330188K24"  
/db\_xref="Ensembl:10090"  
/clone="9330188K24"  
/sex="male"  
/tissue\_type="diencephalon"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
394..2136  
/note="unlabeled protein product; putative  
soluble carrier family 5 (choline transporter), member 7  
(MGDI/MGI:1927126 GB|NM\_022025, evidence: BLASTN, 99%,  
match=1743) "  
/codon\_start=1  
/protein\_id="BAC28702.1"  
/db\_xref="GI:26329927"  
/translation="MSFHEGVATILFLYLIFLVGVAAMKTKNSGNPERSEALIV  
GGRDGLVGGTMTATWVGGYINGTBAVYVPGCGIAMAQAPISLILGLLFF  
AKPMRSKGYVTMLDPFOQYIGKRMGGLPIPLMEEMFMAAIFBALGATTSVITD  
VNISVIALIAILTYLVGLYSAVYTVVQVLCIFIGLWISVPALSHPATDIGFT  
AVHAKYOSPMWLTIESVETWLDNPLMLGIPMOAYFORVLSSSATYAVVISFL  
AAGCLVVALPAICIGAIKASTDMNQTAVGYPDKTEADWILPIVQYLCQVVISF  
FGLGVAASVNSADSSISASMFARNIYOLSPRONSDKESIVWMITVYVFGASA  
TAMALLTKTVVGLWYSSDIYVLIIPOLLCVLFKGNITVGAVAGYVIGLFLRTGG  
EPIYLQPLIFIPGYISDKNGINORFPKTLISMTSPPTNICYSLAKYLPESGTLF  
PKLDFVDAVVAHSEEMDKTILVRENENIKNELAPVAKPROSLTISTFTPEKALLDV  
DSSPBGSGTEEDNQ"

## CDS

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 4306  
Score: 2973.00 Matches: 577  
Percent Similarity: 99.7% Conservative: 1  
Best Local Similarity: 99.5% Mismatches: 2  
Query Match: 99.3% Indels: 0  
Gaps: 0

US-10-724-806-4 (1-580) x AK034415 (1-4306)

QY 1 MetProPheHISValIGluGlyLeuValAlaIleIleLeuPheTyrlLeuLeuIlePheLeu 20  
Db ATGCTTTCCACGTCGAG 453  
QY 21 ValGlyIleTPAlaIleTPAlaIleTPAlaIleTPAlaIleTPAlaIleTPAlaIleTPAla 40  
Db GTTGAATATATGGCGCATGGAACCAAAACAGCGGCAACCCAGAAAGAGCGCAGTGAA 513  
QY 41 AlaIleIleValIGlyIleArgAspIleGlyLeuLeuValIGlyIlePheThrMetThrAla 60  
Db GGCATCATAGTCGGGGCCGTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 573  
QY 514 GGCATCATAGTCGGGGCCGTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 573  
QY 61 ThrTPValIGlyIleGlyIleAsnGlyThrAlaGluAlaValTyrlGlyProGlyCys 80  
Db 574 ACCTGGGTGGAGGAGGCTACATCATGAGCAGCAGGAGAGAGAGAGAGAGAGAGAGAG 633

QY 81 GlyLeuAlaTPAlaHisAlaProIleGlyTyrlSerLeuSerLeuIleLeuGlyIleu 100  
Db GGTCAGCTGGGGCTCAGGACCAATGAGATATCTCGAGCTCAATTTTATAGGTGTG 693  
QY 101 PhePheAlaYSPrometArGSerIleGlyTyrlValThrMetLeuAspProPheIleGln 120  
Db TTTTTCGAAACCTATGCGTTCCAAAGGAGATATGATATGTTTATGACCCATTTCAACAG 753  
QY 121 IleTyrlGlyIleArgMetGlyIleLeuLeuPheIleProAlaLeuMetGlyIleuPhe 140  
Db ATCTATGAAAGCGGATGAGTGGGTGCTTCACTTCCTGACATGATGGAGAGATGTC 813  
QY 141 TPAlaAlaAlaIlePheSerIleAlaLeuGlyIleAlaThrIleSerValIleIleAspAlaAsp 160  
Db TGGGTGAGCAAAATTTTCTGCAATTAGGGGCAACCAACGATGATATGATGATGAT 873  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrlThrLeuValIGlyIle 180  
Db GTGACATATATGCGTATCTCTGACATCTTGGCATCTTTATACCTTATGAGTGG 933  
QY 181 LeuTyrlSerValAlaTyrlThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200  
Db CTCATCTCTGTCGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 993  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
Db ATCACTGTCCTTTTGGCCCTGTCATCTGACATCTGACATCTGACATCTGACATCTG 1053  
QY 221 HisAlaIleTyrlGlnIleSerProTyrlLeuGlyIleThrIleGlnSerValIGlyIleTyrlTrp 240  
Db CATGCTAATATACAGAGTCCCTGCTGGGAAACATGATGATGATGATGATGATGATGAT 1113  
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTyrlGlnAlaTyrlPheGlnArg 260  
Db CTGTATATTTTCTGTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1173  
QY 261 ValLeuSerSerSerSerAlaThrTyrlAlaGlnValIleLeuSerPheLeuAlaIlePheGly 280  
Db GTCCCTCTTATCTCTGACCACTATCTGATGATGATGATGATGATGATGATGATGATGAT 1233  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyIleAlaIleGlyAlaSerThrAspTrp 300  
Db TGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1293  
QY 301 AsnGlnThrAlaTyrlGlyTyrlProAspProIleTyrlGlnGlnAlaAspMetIleLeu 320  
Db AACCAAGCTGCTACGAGGATATCCAGATCCAAAGCTAAGAGAGAGAGAGAGAGATGCTC 1353  
QY 321 ProIleValLeuGlnTyrlLeuCysProValTyrlIleSerPhePheGlyLeuGlyAlaVal 340  
Db CCGATGCTTTCGACATACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1413  
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
Db TCACTGCTGCTCATATCTCTGACATGATGATGATGATGATGATGATGATGATGATGATG 1473  
QY 361 ArgAsnIleTyrlGlnLeuSerPheArgGlnAsnAlaSerAspIleGlnIleValAlaTrpVal 380  
Db CGGAATATCTACACAGCTTCTTCAGCAAAATGATGATGATGATGATGATGATGATGATG 1533  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400  
Db ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1593  
QY 401 ThrValTyrlGlyLeuTyrlTyrlLeuSerSerAspLeuValTyrlIleIleIlePheProGln 420  
Db ACTGCTATGAGCTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1653  
QY 421 LeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyrlAlaValAlaIleGlyTyrlIle 440  
Db CTGCTCTGTATCTTCTTATCATCAAGAGAACCAACTTATGAGGAGAGTGTGATATATTT 1713  
QY 441 PheGlyLeuPheLeuAlaGlyIleThrGlyIleGlnProTyrlLeuTyrlLeuGlnProLeuIle 460

Db 1714 TTGGACTACTCTCGAAGATTACTGGAGGAGACCATATCTACTTGACGCCCTTAATC 1773  
 QY 461 PheTYProGlyTYrTYrSerAspLYsAsnGLYLeuTYrAsnLInArgPheProPheLYs 480  
 Db 1774 TTCTACCTCGGTATTAATCTTGACAGAAATGATATACATACAGAGGTTCCCAATTAA 1833  
 QY 481 ThrLeuSerMetValThrSerPheThrAsnLeuCYsValSerTYrLeuAlaLYsTYr 500  
 Db 1834 ACTCTTCACATGGTTACCTCATCTTTTACCAACATTTGGTTTCTTATCTAGCAAGTAT 1893  
 QY 501 LeuPheGluSerGlyThrLeuProPOLYsLeuAspValPheAspAlaValAlaArg 520  
 Db 1894 CTATTGAAAGTGAACCTTGCTCCAAATATGATGATGATGATGATGATGATGATGATG 1953  
 QY 521 HisSerGluGluAsnMetAspLYsThrLeuValArgAsnGluAsnLeuLYsLeuAsn 540  
 Db 1954 CACAGTGAAGAGAACTGACAGACCATTTAGTCAGAAATGAAATATCAATTAAT 2013  
 QY 541 GluLeuAlaProValLYsProArgLYsSerLeuThrLeuSerSerThrPheThrAsnLYs 560  
 Db 2014 GAACCTTGACACTGTGAACCTCGGACAGCCTTAACCTCACTTCACTTCACTTCACTT 2073  
 QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
 Db 2074 GAGGCCCTCCTTGATGTGATTCAGTCCGAGGGGCTGGAGACTGAAGTAACTTACAA 2133  
 RESULT 3  
 AY413300 1743 bp DNA linear GSS 12-DEC-2003  
 LOCUS Mus musculus HCM4844 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 AY413300  
 VERSION AY413300.1 GI:39769262  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 1743)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,  
 Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene titers  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE  
 AUTHORS 2 (bases 1 to 1743)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B.,  
 Ferrieres,S., Wang,G., Zheng,X.H., White,T.J., Slnsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1.1743  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>1743  
 /locus\_tag="HCM4844"  
 ORIGIN  
 Alignment Scores: 2,33e-255 Length: 1743  
 Score: 2410.00 Matches: 487  
 Percent Similarity: 84.0% Conservative: 0  
 Best Local Similarity: 84.0% Mismatches: 93

Query Match: 80.5% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-724-806-4 (1-580) x AY413300 (1-1743)  
 QY 1 MetProPheHisValGluGlyLeuValAlaAlaLeuPheTYrLeuLeuLeuPheLeu 20  
 Db 1 ATGCTCTTCACGATAGAGAGCTGAGCTATATATCTCTTCACTCTTAATATTCG 60  
 QY 21 ValGlyLeuTyrAlaAlaAlaTyrLYsThrLYsAsnSerGlyAsnProGluGluArgSerGlu 40  
 Db 61 GTTGGAATATGGGCTTCATGAAACCAAAACAGGGGCAACCCAGAGACGACATGAA 120  
 QY 41 AlaAlaLeuValGlyGlyArgAspLeuGlyLeuValGlyGlyPheThrMetThrAla 60  
 Db 121 GCCATATAGTGGGGGCGGTGACATGGTTGTTGGTGGTGTTCATCATGACAGNN 180  
 QY 61 ThrTyrValGlyGlyGlyTYrLeuAsnGlyThrAlaGluAlaValTYrGlyProGlyCYs 80  
 Db 181 NNN 240  
 QY 81 GlyLeuAlaTyrAlaHisAlaProLeuGlyTYrSerLeuSerLeuLeuGlyGlyLeu 100  
 Db 241 NNN 300  
 QY 101 PhePheAlaLYsProMetArgSerLYsGlyTYrValThrMetLeuAspProPheLYsGln 120  
 Db 301 NNN 360  
 QY 121 IleTYrGlyLYsArgMetGlyGlyLeuPheIleProAlaLeuMetGlyLeuMetPhe 140  
 Db 361 NNN 420  
 QY 141 TyrAlaAlaAlaAlaPheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
 Db 421 NNN 480  
 QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTYrThrLeuValGlyGly 180  
 Db 481 GTGAACATATCGGTCACTTGTCTCTGACATCATTTGCCATTTTATACCCAGTGGGG 540  
 QY 181 LeuTYrSerValAlaTYrThrAspValGlnLeuPheCYsIlePheIleGlyLeuTyr 200  
 Db 541 CTCTACTCTGTGGCATATCTGATGTTGTTCAGCTATTCGCAATTTTAAAGACCTGTG 600  
 QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
 Db 601 ATCAGTGTCCCTTTTCCCTGTCACATCTCGACGTACCGACATCGGATTCACAGCTGTG 660  
 QY 221 HisAlaLYsTYrGlnSerProTyrPLeuGlyThrIleGlySerValGluValTYrThrTyr 240  
 Db 661 CATGCTAAATCCAGAGTCCCTGCGGAGAACCATTTGAATCAAGCTACACTGG 720  
 QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleLeuProTyrGlnAlaTYrPheGlnArg 260  
 Db 721 CTGTATATATTTCTGTATTATGATGCTGGGTGAATCCCATGCAAGCTCTACTCCAGAGG 780  
 QY 261 ValLeuSerSerSerSerAlaThrTYrAlaGlnValLeuSerPheLeuAlaAlaPheGly 280  
 Db 781 GTCTCTCTTCACACCCACGACCATTAATGCTCAGAGTCTGCTCTCGGACGCTTTGGG 840  
 QY 281 CysLeuValMetAlaLeuProAlaIleCYsIleGlyAlaIleGlyAlaSerThrAspTyr 300  
 Db 841 TGCCTGTATGCTCTTACCCGCAATATGATATGAGAGCTATTTGAGGTTTCAACAGCTGG 900  
 QY 301 AsnGlnThrAlaTYrGlyTYrProAspProLYsThrLYsGluGluAlaAspMetCileu 320  
 Db 901 AACCAAGCTGCCATACGGGTATCCAGATCCCAAGACTAAGAGAGACAGACATGATTC 960  
 QY 321 ProIleValLeuGlnTYrLeuCYsProValTYrIleSerPhePheGlyLeuGlyAlaVal 340  
 Db 961 CCGATGTTCTGACGATCCTGCTGCTGTGATCATCTCTTCTTTGGGCTTGCTGCTTT 1020

[illegible]

TITLE	Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sniinsky, J.J., Adams, M.D. and Cargill, M.
JOURNAL	Direct Submision
COMMENT	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..1743
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>1743 /locus_tag="HCM4844"
ORIGIN	
Alignment Scores:	
Prod. No.:	2..13e-238
Score:	2257.00
Percent Similarity:	81.4%
Best Local Similarity:	77.6%
Query Match:	75.4%
DB:	14 Gaps: 0
US-10-724-806-4 (1-580) x AY413298 (1-1743)	
QY	1 MetProPheHisValGluGlyLeuValAlaIleIleuPheTyrLeuIlePheLeu 20
DB	1 ATGGCTTTCATGATGGAAGACTGATAGTATCATCGTGTTCTACCTTCTATTTCCTG 60
QY	21 ValGlyIleTPrpAlaAlaTPrpLysThrIysAsnSerGlyAsnProGluGluArgSerGlu 40
DB	61 GTTGAAGATATGGGCTCGTGGAGAAACCAAAAAGTGGCGACGCGAAGAGCCGCGAA 120
QY	41 AlaIleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPheThrMetThraAla 60
DB	121 GCCATCATAGTTGGTGGCCGAGATATGGTTATTTAGTTGGATTTACCATGACAGNN 180
QY	61 ThrTrpValGlyGlyGlyTyrIleAsnGlyThraIleGluAlaValTyrGlyProGlyCys 80
DB	181 NNN 240
QY	81 GlyLeuAlaTPrpAlaHisAlaProIleGlyTyrSerLeuSerIleLeuGlyGlyLeu 100
DB	241 NNN 300
QY	101 PhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheGlyGln 120
DB	301 NNN 360
QY	121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB	361 NNN 420
QY	141 TyrPheAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB	421 NNN 480
QY	161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB	481 ATGCACATTTCTGTGATCATCTCTGCACATTTGCCACTGTGTACACACTGTCGGAGGG 540
QY	181 LeuTyrSerValAlaIleTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
DB	541 CTCATATTCGTGGCCACACTGATGTGCTGACGCTCTTTTCATTTTTGAAGGCGTGG 600
QY	201 IleSerValProPheHisAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
DB	601 ATCAGCGTCCCTTTCGATTCATCATCTCGTAGTCGAGCGACACATCGGGTTCACTGCTG 660
QY	221 HisAlaLysTyrGlnSerProTPrpLeuGlyThrIleGluSerValGluValTyrThrTPrp 240
DB	661 CAGGCCAAATACCAAAAGCCGTGGCTGGGAACTGTTGATCATCATCTGAAGCTTACTCTGG 720

```

QY 241 LeuAspAsnPhenLeuLeuMetLeuGlyGlyIleProTrogInaIaTyPheGlnArg 260
Db 721 CTTGATAGTTTCTGTTGTTGATGCTGGGATGATCCATGGACAGATACCTTTCAGAG 780
QY 261 ValLeuSerSerSerSerAlaThrTyAlaGlnValLeuSerPheLeuAlaIaPheGly 280
Db 781 GTTCTCTCTCTCTCCAGCAGCCTTATGCTCAAGTCTGCTGCTTCCGAGAGCTTTCGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTTP 300
Db 841 TGGCTGGTATGGCTGATCCAGCAGCCTTATGCTGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 301 ArgGlnThrAlaTyGlyTyProAspProLysThrGlyGlnIaIaSpMetIleLeu 320
Db 901 AACCAAGCTGATATGGCTTCCAGATCCAGAGATACAGAGAGGAGAGAGATGATTTTA 960
QY 321 ProIleValLeuGlnTyLeuCysProValTyIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CCAATTGTTCTGAGATGCTGCTGCTGCTGATATTTCTTCTTGGTGGTGGAGTT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360
Db 1021 TCGCTGCTGTTATGATCATCAGACAGATCTTCCATCTTGCAGCAAGTCCAGATTTGGA 1080
QY 361 ArgAsnIleTyGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTTPVal 380
Db 1081 CGGAACATCATCAGCCTTCTTCCAGCAAAATGCTTCCAGCAAAAGAAATCGTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGCGATGATCAGATGTTGTTGTTGAGATGCTGCAACAGCCAGCTTGCAGAGAAA 1200
QY 401 ThrValTyGlyLeuTyPheLeuSerSerAspLeuValTyIleIleIlePheProGln 420
Db 1201 ACTGTGATATGGCTGCTGATCCTCAGTTCGACCTTGTATCATGTTCTTCCCCCAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyGlyAlaValaIaGlyTyIle 440
Db 1261 CTGCTTGTGATCTTCTTGTATGAGGAGCAACCACTTATGGGCGCTGGCAGGTTATGTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyLeuTyLeuGlnProLeuIle 460
Db 1321 TCGGCTCTTCTCTGAGATTAACAGAGGAGGAGCAGATCTGATCTTCCAGCCCTTGATC 1380
QY 461 PheTyPProGlyTyTyTySerAspLysAsnGlyIleTyAsnGlnArgPheProPheLys 480
Db 1381 TTCTACCTGGCTATTAACCTGATGATTAATGATATATATATATATATATATATATATAA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyLeuAlaLysTy 500
Db 1441 ACACTTGCATGATGATCATCATCTTAACCAACATTTGATCTTATGAGCAAGTAT 1500
QY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValaIaArg 520
Db 1501 CTAATTTGAAGTGAAGCTTGCACCACTTAATTAATGATGATTAATGATGATGATGATG 1560
QY 521 HisSerGlyGlnAsnMetAspLysThrIleLeuValArgAsnGlnIleLysLeuAsn 540
Db 1561 CACAGTGAAGAAACATGATTAAGACATTTCTGTCAAAAATGATAATATAATTAATTAAT 1620
QY 541 GlnLeuAlaProAlaLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACCTTGCATTTGAGGAGCAGACAGAGATGACCTTGCAGCTCACTTCAACATTA 1680
QY 561 GlnAlaIleLeuLeuAspValaIaAspSerSerProGlyGlySerGlyTyThrGlnAspAsnLeuGln 580
Db 1681 GAGGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740

```

RESULT 5  
 AY413299 1743 bp DNA linear GSS 12-DEC-2003  
 LOCUS Pan troglodytes HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,  
 DEFINITION

```

ACCESSION      Genomic survey sequence.
VERSION        AY413299.1 GI:39769261
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
               Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homiidae; Pan.
REFERENCE     1 (bases 1 to 1743)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
               Adams,M.D. and Cargill,M.
TITLE        Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene clusters
JOURNAL       Science 302 (5652), 1960-1963 (2003)
PUBMED       14671302
REFERENCE     2 (bases 1 to 1743)
AUTHORS      Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,J.J.,
               Adams,M.D. and Cargill,M.
TITLE        Direct Submission
JOURNAL       Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
COMMENT      This sequence was made by sequencing genomic exons and ordering
               them based on alignment.
FEATURES      Location/Qualifiers
               source
                 1..1743
                 /organism="Pan troglodytes"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9598"
                 <1..>1743
                 /locus_tag="HCM4844"
ORIGIN
Alignment Scores:
Pred. No.:    9,85e-238      Length:    1743
Score:        2251.00       Matches:    450
Percent Similarity: 81.2%   Conservative: 21
Best Local Similarity: 77.6% Mismatches:    109
Query Match:  75.2%        Indels:      0
DB:           14           Gaps:      0
US-10-724-806-4 (1-580) x AY413299 (1-1743)
QY 1 MetProPheHisValGlyLeuValaIaIleLeuPheTyPheLeuIlePheLeu 20
Db 1 ATGCGCTTCCATGATGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 60
QY 21 ValGlyIleTTPAlaIaIaTTPlySThrLysAsnSerGlyAsnProGlyGlnArgSerGln 40
Db 61 GTTGGATATATGGCTGCTGCGAGAACCAAAACAGTGGACGCCCAAGAGCGCACGAA 120
QY 41 AlaIleIleValaIaGlyLysArgAspIleGlyLeuLeuValaIaGlyPheThrMetTTPAla 60
Db 121 GCCATCATATGTTGGTGGCGAGATATTTGGTTTATTTGGTGGTGGATTTACCATGACANN 180
QY 61 ThrTTPValaIaGlyGlyTyIleAsnGlyThrAlaGlnAlaValaTyGlyProGlyCys 80
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 81 GlyLeuAlaIaTTPAlaHisAlaProIleGlyTySerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 101 PhePheAlaLysProMetArgSerLysGlyTyValaIaThrMetLeuAspProPheLysGln 120
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY 121 IleTyGlyLysArgMetGlyGlyLeuLeuPheIleProAlaIleMetGlyGlnMetPhe 140

```







CONSRMT	Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.
CONSRMT	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
PUBMED	11217851
REFERENCE	5
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, U., Bono, H., Kondo, S., Nikaido, I., Oساتو, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baladrelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batelov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusch, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dregant, T. A., Fletcher, C. F., Forrest, A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzielski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, C., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, K., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Itokawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
CONSRMT	FANTOM Consortium
TITLE	Analysis of the mouse transcriptome based on functional annotation
JOURNAL	of 60,770 full-length cDNAs
PUBMED	Nature 420 (6915), 563-573 (2002)
REFERENCE	6
AUTHORS	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fitch, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzus, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., Ambesi-Impombato, A., Apweiler, R., Attalaya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Belsel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chin, K. P., Choudhary, V., Christofideis, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeno, K., Iwama, A., Ishikawa, T., Jaki, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsø, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S. P., Kruger, A. E., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, P., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Misgurne, F., Miyake, S., Morits, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K. C., Pavan, W. J., Pavoni, G., Pesole, G., Petrovsky, N., Plaza, S., Reed, J. F., Reid, J. F., Ring, B. Z., Rindwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schombach, C., Sekiguchi, K., Semple, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalati, B., Sperling, S., Scupka, E., Sugiyara, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamashita, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bull, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusch, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M.,
CONSRMT	Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Komno, H., Nakano, K., Niimura, T., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanuki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
CONSRMT	FANTOM Consortium
TITLE	The transcriptional landscape of the mammalian genome
JOURNAL	Science 309 (5740), 1559-1563 (2005)
PUBMED	16141072
REFERENCE	7
AUTHORS	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Fitch, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batelov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
CONSRMT	RIKEN Genome Exploration Research Group
TITLE	Antisense transcription in the mammalian transcriptome
JOURNAL	Science 309 (5740), 1564-1566 (2005)
PUBMED	16141073
REFERENCE	8 (bases 1 to 4122)
AUTHORS	Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Horii, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Komno, H., Murata, M., Nakamura, M., Niimura, Y., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y.
CONSRMT	Direct Submission
TITLE	Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-reegsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/
FEATURES	Location/Qualifiers
source	1..4122
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM DB:C530033B06"
	/db_xref="taxon:10090"
	/clone="C530033B06"
	/tissue_type="spinal cord"
	/clone_1lb="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="12 days embryo"
	<279..1160
	/note="unnamed protein product; putative
	solute carrier family 5 (cholesterol transporter), member 7
	(MGD MG1:1927126 GB NM_022023, evidence: BLASTN, 99%,
	match=2969)
	start codon is not identified"
	/codon_start=1
	/protein_id="BAE37635.1"
	/db_xref="GI:74211081"
	/translation="ALSCGFYCSQDMNTAYGYDPKTKREADMILPIVLYLCVY
	ISFGLGSAVMSADSSIISSASFRANKIYQSFQNSADKEIVWNRITLVFG
	ASATMALTLKTVYELWYISDLVYIIIFPLLCVLFKGTNYGVAVAYIGFLFRI
	TGGEYLYIOPILFPGYYSKDKGKLYNQGPFKPTLSMTVPTNLCVSIAYLVPSG
	LDVPSDPSGSGEDMLQ"
	4100..4105
	/note="putative"
	4122
	polyA_signal
	polyA_site

## Alignment Scores:

Pred. No.: 3,956-149 Length: 4122  
 Score: 1457.00 Matches: 255  
 Percent Similarity: 97.6% Conservative: 1  
 Best Local Similarity: 97.3% Mismatches: 7  
 Query Match: 48.7% Indels: 0  
 DB: 6 Gaps: 0

US-10-724-806-4 (1-580) x AK164116 (1-4122)

```

QY 288 AAlaieCySileGAlAlaAlaieGlyAlaSerThrAspTrpAsnGlnThrAlaTyrGlyTyr 307
DB 279 GGTCTCAAGTTGTGGTGTTCTTCTGCTCCACAGACTGGAAACCAAGCTCCACGGGTAT 338
QY 308 ProAspProLysThrLysGlnGlnAlaAspMetIleLeuProIleValIleuGlnTyrLeu 327
DB 339 CCAGATCCCAACACTRAGAGAGAGACAGACATCTCCGATCGTTCTGCAGTACCTC 398
QY 328 CysProValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSer 347
DB 399 TGCCCTGCTACATCTCTTCTTGGGCTTGGCTGCTTTCAGCTGCTGCATGCTCCA 458
QY 348 AlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeuSer 367
DB 459 GCTGACCTGCTCACTCTGCTGGCGAGTTCTATGTTTGTCTGGAAATTTCTACAGCTTTC 518
QY 368 PheArgGlnAsnAlaSerAspLysGlnIleValITrpValMetArgIleThrValIleuVal 387
DB 519 TTCAGACAAATGATGCATGACAGCAAGAAATTTGTGGGTCAATGAGATCATCTGGCTGTGG 578
QY 388 PheGlyAlaSerAlaThrAlaMetAlaLeuLeuTyrIleValTyrGlyLeuTyrTyr 407
DB 579 TTCGAGAGATCTGCACACGCCATGCTTGTGCAGAACACGCTGTATGGCTCTGTGAC 638
QY 408 LeuSerSerAspLeuValTyrIleIleIlePheProGlnIleLeuCysValIleuPheIle 427
DB 639 CTGAGCTCTGACCTGCTCTACATCATCTTCCACAGCTCTCTGTATCTCTTCACT 698
QY 428 LysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIle 447
DB 699 AAAGAACCAACACTTATGGGGGAGTGTGCTTATATTTTGGACTATTTCTCGAAGATT 758
QY 448 ThrGlyGlyGlnProTyrLeuTyrIleuGlnProIleuIlePheTyrProGlyTyrTyrSer 467
DB 759 ACTGAGAGAGAGACCATATCTATCTATCTGACGCCCTTAATCTTACCTGCTGTATTA 818
QY 468 AspLysAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThrSer 487
DB 819 GACAAAGATGTATATACATCATCAGAGTTCCCAATTAACCTCTCTCCATGGTTACCTCA 878
QY 488 PhePheThrAsnIleCysValSerTyrLeuAlaValTyrLeuPheGlnuSerGlyThrLeu 507
DB 879 TTCCTTACCAACATTTGTTGTTCTTATCTAGCCCAAGTATCTATTGGAAGTGGAACTTGG 938
QY 508 ProProLysLeuAspValPheAspAlaValIleAlaArgHisSerGlyLeuAsnMetAsp 527
DB 939 CCTCAAAATTAAGATGATTTGATGCTGTGTGCGAAGGCAAGTGAAGAAACATATGAC 998
QY 528 LysThrIleLeuValArgAsnGlnuAsnIleLysLeuAsnGlnIleuAlaProValLysPro 547
DB 999 AAGACCATTTCTAGTCAAGAAATGAATTAATCAAAATTAATGAACCTTGCAACCTTGA 1058
QY 548 ArgGlnSerLeuThrLeuSerSerThrPheThrAsnLysGlnAlaLeuLeuAspValAsp 567
DB 1059 CGGCGAGAGCCCTAACCCCTCAGTTCAACTTTCACCAATTAAGAGAGCCCTCTGATGTGAT 1118
QY 568 SerSerProGlnGlySerGlyThrGlnuAsnLeuGln 580
DB 1119 TTCAGTCCGAGAGGGGTCTGGGACTGAAGATTAACCTTAACA 1157

```

RESULT 7  
 , CK770440/c

```

LOCUS CK770440 882 bp mRNA linear EST 20-FEB-2004
DEFINITION 958613 MARC 1BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CK770440
VERSION CK770440.1 GI:42724534
KEYWORDS EST
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 882)
REFERENCE Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
AUTHORS Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Perteaux,G., Holt,I., Karameyeva,S., Liang,F.,
Quckenbush,J. and Keefe,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
JOURNAL libraries and construction of a gene index for cattle
PUBMED Genome Res. 11 (4), 626-630 (2001)
11282978
COMMENT Contact: Smith RPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: 95 row: L column: 20
Seq primer: CTAAATACGACCTCATATAGG.
FEATURES
Location/Qualifiers
source 1..882
/mol_type="mRNA"
/organism="Bos taurus"
/db_xref="taxon:9913"
/lisue_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 1BOV"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
ORIGIN
Alignment Scores:
Pred. No.: 1,886-126 Length: 882
Score: 1243.00 Matches: 241
Percent Similarity: 96.6% Conservative: 13
Best Local Similarity: 91.6% Mismatches: 9
Query Match: 41.5% Indels: 0
DB: 5 Gaps: 0
US-10-724-806-4 (1-580) x CK770440 (1-882)
QY 318 MetIleuProIleValIleuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeu 337
DB 880 ATGATCTTGGCGATTTGCTCAAGTATCTGCCCCGATGACATTTTCTTATTTGCTTT 821
QY 338 GlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSer 357
DB 820 GGAGCGCTTTCGCTGCGTGTATGCTCTCAGAGATTTCTTCATCTTGTGCAGCAAGTTCG 761
QY 358 MetPheAlaArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIle 377
DB 760 ATGTTGCTCGCAACATCTACAGCTTTCATTCACAAATGCTTTCGACAAAGAGATA 701
QY 378 ValITrpValMetArgIleThrValIleuValPheGlyAlaSerAlaThrAlaMetAlaLeu 397
DB 700 GTCTGGGTCATGCGGATCAAGGATTTTGTGAGCTTCTGCGATGACCATGAGCTTGG 641
QY 398 LeuThrLysThrValTyrGlyLeuTyrPyrLeuSerSerAspLeuValTyrIleIleIle 417
DB 640 CTAAAGAAAGAGGTATAGGGGCTGTGATCTGACCTTGACCTGTGATCATCATCATC 581

```

```

OY 418 PheProGlnLeuLeuCyValLeuPheIleLeuGlyThrAsnThrTyrGlnAlaValAla 437
DB 580 TTTCCGCAATTGCTGCTGCTGCTCTTTCATCAGAGGATCCACACGATGAGGGCGGGCA 521
OY 438 GtYrYrIlePheGlyLeuPheLeuArgIleThrGlyGlyValuProTyrLeuTyrLeuGln 457
DB 520 GGTACATCTCGGGGCTTTTCCGCGAGACACTGGGGGAGGAGCTTACCTGAACTGCAG 461
OY 458 ProLeuIlePheTyrProGlyTyrTyrTyrSerAspIleValIleTyrAsnGlnAlaPhe 477
DB 460 CTTTGATCTTTTATCCCTGCTATTAATGATTAAGGTGGCATATATTAACCGAGATTTC 401
OY 478 ProPheIleThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeu 497
DB 400 CATTTTAAACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 341
OY 498 AlaIleTyrLeuPheGlyLeuSerGlyThrLeuProProIleuAspValPheAspAlaVal 517
DB 340 GCCAATATACCTGTTGAAAGTGAACCTTACACCAAAATTTGATGATTCGATCTGTT 281
OY 518 ValAlaIleArgHisSerGlnGluAsnMetAspIleThrIleLeuValArgAsnGlnAlaIle 537
DB 280 GTTCGCAAGGACACAGTGAAGAAAACATGACACACACATTCCTGTTGAAATGAAACATT 221
OY 538 LysLeuAsnGlnLeuAlaProValIleProArgGlnSerLeuThrLeuSerSerThrPhe 557
DB 220 AAATCTGATGAACTCCGACCTGTGAAGCCACGACAGCATTAATCTTCAGCTGACCTTC 161
OY 558 ThrAsnLysGlnAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAsp 577
DB 160 ACCATCATGAGAGGCCCTCTGGAATGATGATTCAATGACAGAGGCTCTGGAGACTGAGAT 101
OY 578 AsnLeuGln 580
DB 100 AATTATACAG 92

RESULT 8
LOCUS CD350164 707 bp mRNA linear EST 08-JUL-2003
DEFINITION UT-M-FY0-cel-h-10-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6851099 5', mRNA sequence.
ACCESSION CD350164
VERSION CD350164.1 GI:31141679
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 707)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouseefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seg primer: pyx-5,
Location/Qualifiers
1. 707
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6851099"
/tissue_type="whole brain"

```

```

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_FY0"
/note="Organ: Brain; Vector: pyx-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Herman Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 7,21e-107 Length: 707
Score: 1065.00 Matches: 219
Percent Similarity: 96.9% Conservative: 1
Best Local Similarity: 96.5% Mismatches: 2
Query Match: 35.6% Indels: 5
DB: Gaps: 0

US-10-724-806-4 (1-580) x CD350164 (1-707)

OY 1 MetProPheHisValGlyLeuValAlaIleLeuPheTyrLeuLeuIlePheLeu 20
DB 27 ATGCTCTTCCAGTAAAGAGAGCTGAGTATATATCTCTTACCTCTTATATTTCTG 86
OY 21 ValGlyIleTyrAlaIleTyrIleAsnSerGlyAsnProGluGluArgSerGln 40
DB 87 GTTGAATATGGCTGCATGAAACCAAAACAGGGCAACCCAGAGCGCAGTAA 146
OY 41 AlaIleIleValGlyValArgAspIleGlyLeuLeuValGlyIlePheThrMetTrpAla 60
DB 147 GCCATCATGCTGGGGGCCGCTGACATTTGCTTGTGTGGTGTGTTTACATGACAGCC 206
OY 61 ThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGlnAlaValTyrGlyProGlyCys 80
DB 207 ACCTGGTTTGAAGAGCTACATCATGAGGACAGCAAGACTGATGGCCAGGTGT 266
OY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeu 100
DB 267 GGTCTGCTGGGCTCAGGACCCCATTTGATATCTCTGATGCTTAATTTAGGTGCTG 326
OY 101 PhePheAlaIleProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGln 120
DB 327 TTTTTCGAAACCTATGCGTTCCAAAGGATATGATGATGATGATGATGATGATGATG 386
OY 121 IleTyrGlyLysArgMetGlyGlyLeuPheIleProAlaLeuMetGly-GluMetPhe 140
DB 387 ATCTATGAAAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446
OY 140 eTyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIle-SerValIleIle-AspVal 159
DB 447 CTGGGCTGACGCAATTTTCTGCACTTAGGGGCGACCATTCAGCGATCATTTGGATGTG 506
OY 160 AspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGly 179
DB 507 GATGTAAACATATCGCTCATTTGCTCTGCACTCATTTGCCATTTTATACCTAGTGGT 566
OY 180 GlyLeuTyrSerValAlaTyrThr-AspValAlaGlnLeuPheCysIlePheIleGlyLe 199
DB 567 GGGCTTACTCTGTGTCATATCTGATGTGTCAGCTATTTCTGATTTTATAGACT 626
OY 199 UTTP-IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThr 219
DB 627 GTGGATCATGTCCTTTTGGCTTCTGACATCTGCAAGTACGACGACATGATGATCAG 686

```

QY 219 lavalhiala 222  
 Db 687 CTGTGCATGCT 697

RESULT 9  
 CT403380  
 LOCUS Sus scrofa genomic clone CH242-157K9, genomic survey sequence.  
 DEFINITION  
 ACCESSION CT403380  
 VERSION CT403380.1 GI:80060743  
 KEYWORDS GSS.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae; Suidae; Sus.

REFERENCE  
 1 (bases 1 to 828)  
 HUMPHRAY, S.J., PLUMB, R.W. and DURHAM, J.L.  
 Direct Submission  
 Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished  
 This sequence was generated from the SP6 end of BAC 157K9. 157K9 is part of the CHORI-242 BAC Library created by P. de Jong. Further details: [http://www.sanger.ac.uk/Projects/S\\_scrofa/](http://www.sanger.ac.uk/Projects/S_scrofa/).

FEATURES  
 source  
 1..828  
 /organism="Sus scrofa"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9823"  
 /clone="CH242-157K9"  
 /tissue\_type="white blood cells"  
 /note="vector pTARBAC1.3\_BamHI sex female"

ORIGIN

Alignment Scores:  
 Pred. No.: 5,336-100 Length: 828  
 Score: 1004.00 Matches: 201  
 Percent Similarity: 85.34 Conservative: 19  
 Best Local Similarity: 77.94 Mismatches: 32  
 Query Match: 33.54 Indels: 6  
 DB: 14 Gaps: 2

US-10-724-806-4 (1-580) x CT403380 (1-828)

QY 327 leucysprovaltyrileserphephegyleuglyalavalaseralavalmetser 346  
 Db 9 GTCTGTCATGTTCGTCGACACTTTATGCTCTCATTTTATGCGTATTTTCATG 68  
 QY 347 SerAlaaspserSerile-----leuSerAlaSerSerMetpheAlaarg 361  
 Db 69 GTCCCTTAGAGTAGCTTTAAACACAGAGAGCATCTCATGCTCAGCCATCTGAGGAA 128  
 QY 362 AsnileTyrglnleuSerPhehArglnAsnAlaSerAspLysGlnileValTnpValmet 381  
 Db 129 AACTGACTCGATGCGATTTCTTGAG--GCTTCGACAGGAGATCGCTGCGGCATG 185  
 QY 382 ArgileThrValIleuValPheGlyAlaSerAlaThrAlaMetAlaIleuendThryrThr 401  
 Db 186 CGGATCACAGATTGTGTGTTGTCGCTGTGCAACGCGATGCTGTGTCACCAAGACC 245  
 QY 402 ValTyrglyLeuTyrglyLeuSerSerAspLeuValTyrlleIlelePheProGlnleu 421  
 Db 246 GTGTATGGGCTCTGTGACTCTCGACCTCGACCTGTATCATATATCTTCCCGCAGCTG 305  
 QY 422 LeuCyValIleuPheIlelysglyThrAsnThryrglyAlaValAlaGlyTyrllePhe 441  
 Db 306 CTCTGTGCTCTTATCATCAAGGAGCAACACAGTACGGGGCCGTGCGACGGTACATGCT 365  
 QY 442 GlyLeuPheLeuArgIleThrIlglyGlnupProTyrlleTyrlleGlnProleuIlePhe 461

Db 366 GGCCTTTTCTGAGGAGTACCGGTGAGAGCCATACCTGAACCTCGACCCCTTATCTTT 425  
 QY 462 TyrProglyTyrglySerAspLysAsnGlyIleTyraGlnIaArgPheProPheIleThr 481  
 Db 426 TACCTGGTATATACGTTGAAAAAATGATATATATATACAGAGATCCCATTTAAACC 485  
 QY 482 LeuSerMetValThrSerPheThrAsnIleCyValSerTyrlleAlaIleTyrlleu 501  
 Db 486 CTTCGCAATGCTTCACTCTTCTTATCCACATTTGCACTCTTATGCGCAAAATATCTA 545  
 QY 502 PheGlnSerGlyThrLeuProPheProLysLeuAspValPheAspAlaValAlaArgHis 521  
 Db 546 TTTGAAAGTGAACCTTCCACCAAAATTAAGATATGTTGATGCTGTGTTCCAGACAC 605  
 QY 522 SerGlnIuAsnMetAspLysThrIleLeuValArgAsnGlnAsnIleLysLeuAsnGlu 541  
 Db 606 AGTGAAGAAACATGATATAGACCATTCCTGTCGCAAAATGAAAAACATTAAATTAACGAA 665  
 QY 542 LeuAlaProValIyProArgIlnSerleuThrLeuSerSerThrPheThrAsnIleGlu 561  
 Db 666 CTTCGACCTGTGAACCTTCGGAGGCTTAACCTGACCTCGACTTCACCATCAGAGAG 725  
 QY 562 AlaLeuAspValAspSerSerProGlnIySerGlyThrGluAspAsnLeu 579  
 Db 726 GCCTTCTCAACATGATTTCCAGCCCAAGAGTTCTGGGACAGAAACATTTTG 779

RESULT 10  
 Dw665738  
 LOCUS Dw665738 988 bp mRNA linear EST 19-JAN-2006  
 DEFINITION CNB333-A09.yid-s SHGC-CNB2 Gasterosteus aculeatus cDNA clone  
 CNB333-A09 5', mRNA sequence.  
 DW665738  
 DW665738.1 GI:85459051  
 EST.  
 Gasterosteus aculeatus (three spined stickleback)  
 ORGANISM Gasterosteus aculeatus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.  
 1 (bases 1 to 988)  
 Kingsley, D.M., Peichel, C., Knecht, A., Balabhadr, S., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.  
 Expressed sequence tags from Gasterosteus aculeatus (2004)  
 Unpublished (2004)  
 Contact: Grimwood, Jane  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 975 S California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 320 5917  
 Fax: 650 320 5801  
 Email: jane@shgc.stanford.edu  
 Plate: 333  
 High quality sequence stop: 861.  
 Location/Qualifiers  
 1..988  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="mRNA"  
 /strain="Conner Creek sticklebacks, WA USA"  
 /db\_xref="taxon:69293"  
 /clone="CNB333-A09"  
 /sex="mixed male and female"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="SHGC-CNB2"  
 /note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTTCTAGATCGCAGCGCGGCC(T)25-3'). Following

second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5' prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site), and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: [http://www.openbiosystems.com/cdna\\_library\\_construction\\_fa](http://www.openbiosystems.com/cdna_library_construction_fa) g.phb8 The primary library was transformed and amplified in DH10B (TI phage resistant) bacteria."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,67e-98 Length: 988  
Score: 988.50 Matches: 192  
Percent Similarity: 85.2% Conservative: 26  
Best Local Similarity: 75.0% Mismatches: 35  
Query Match: 33.0% Indels: 4  
Gaps: 1

US-10-724-806-4 (1-580) x DW655738 (1-988)

QY 1 MetProPheHisValGluGlyLeuValAlaIleLeuPheTyrLeuLeuIlePheLeu 20  
DB 223 ATGACCATCATATAGAGGGGCTTGCTATCGCATCTTCTATCTAGATCGTGTGC 282  
QY 21 ValGlyIleThrAlaIleTyrIleValAsnSerGlyAsnProGlu-----GluArg 38  
DB 283 GTGGGATCGTCGAGGGGCGATGAAAAAACAACCTCCGGGAGGGGAGGGGACCGCCG 342  
QY 39 SerGluAlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheTyrMet 58  
DB 343 ACGGAACCATCATATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 402  
QY 59 ThrAlaThrTyrValGlyGlyGlyTyrIleAsnGlyTyrThrAlaGluAlaValTyrGlyPro 78  
DB 403 ACAGGACGCTGGGTCGAGAGAGATATATATATATATATATATATATATATATATAT 462  
QY 79 GlyCysGlyLeuAlaThrAlaHisAlaProIleGlyTyrSerLeuLeuLeuLeuGly 98  
DB 463 GCGTACGTTGGCTTGGCTTCAAGCTCCCTTGGATATGACCTGCTGTTGTGGGG 522  
QY 99 GlyLeuPhePheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPhe 118  
DB 523 GGTCTATTTTTCGCCAAGCCCATCGCTCAGCGGTTAGCTCAACATGTTGGACCGCTTC 582  
QY 119 LysGlnIleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138  
DB 583 CAGCAGATCATCGGGAACGATGCGGTGCTCTTCTATCTAGATCGTCACTGAGAGAG 642  
QY 139 MetPheThrAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158  
DB 643 ATCTTCTGTCGCGCCCATCTTATCCGCTCGGTCTGATCTGATGTCATCGTGGAGAC 702  
QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178  
DB 703 ATCAACATTAAGATGTCGCTGATCTAGCGCTCATTCGATTTTACACCTTGCTT 762  
QY 179 GlyGlyLeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGly 198  
DB 763 GAGGAGCTTATCTCGGTGCTCAGCGATGTCGTCAGCTCTTGCATCTTCTTGGC 822  
QY 199 LeuThrIleSerValProPheAlaLeuSerHisProAlaValAlaAspIleGlyPheThr 218  
DB 823 CTGTGATACAGGCTCCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 882  
QY 219 AlaValHisAlaLysTyrGlnSerProTyrPheGlyThrIleGlnSerValGluValTyr 238  
DB 883 GCGATGAGAGAGGTGATACGATGCGCTGAGAGGAGCATCCACAAGGAGACAC--CG 940  
QY 239 ThrTyrLeuAspAsnPhe-LeuLeuLeuMetLeuGlyIlePro 253

DB 941 GTGTGATCATATACCTTCTGCTCCCTCTGATGCTGAGNAGATTTTCCC 986

## RESULT 11

BY272598

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

## REFERENCE

AUTHORS

BY272598 669 bp mRNA linear EST 17-DEC-2002  
BY272598 RIKEN full-length enriched, 6 days neonate medulla  
oblongata Mus musculus cDNA clone B730003H24 5', mRNA sequence.  
BY272598  
BY272598.1 GI:27140725  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Mus.  
1 (bases 1 to 669)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,  
Chotina, C., Corbani, U.E., Cousins, S., Dalla, E., Dregant, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, D.U., Q.D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K.,  
Sulana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishi, Y.,  
Iton, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9226  
Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,  
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,  
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Substitution  
Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

## FEATURES

source

Location/Qualifiers

1..669

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="B73003H24"

/tissue\_type="medulla oblongata"

/dev\_stage="6 days neonate"

/lab\_host="DH10B"

/clone\_lib="RIKEN full-length enriched, 6 days neonate  
medulla oblongata"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5',  
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTVA 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 20.0 and subtraction to Rot = 459.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGATCTCCAGTATTAATTAATCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC I."

## ORIGIN

## Alignment Scores:

Pred. No.: 2..93e-98 Length: 669  
 Score: 987.00 Matches: 192  
 Percent Similarity: 99.0% Conservative: 1  
 Best Local Similarity: 98.5% Mismatches: 2  
 Query Match: 35.0% Indels: 0  
 DB: 4 Gaps: 0

US-10-724-806-4 (1-580) x BY727598 (1-669)

OY 1 MetProPheHisValGlnGlyLeuValAlaIleIleuPheTyrIleuLeuPheLeu 20  
 DB 81 ATGTCTTCCAGTGAAGAGACTGTAGCTATTATTCCTTTACCTCTTAAATTTCTG 140  
 OY 21 ValGlyIleTyrPAlaIleTyrIleuValAsnSerGlyAsnProGluGluArgSerGlu 40  
 DB 141 GTTGAATATATGGGCTCCATGAGAAACCAAAACAGGGGCAACCCAGAAAGCGCAGTGA 200  
 OY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
 DB 201 GCCATCATATGTCGGGGCCGTGACATTTGTTGGTGGTGGTTTACCATGACAGCC 260  
 OY 61 ThrTTPValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyGly 80  
 DB 261 ACCTGGGTGGAGAGGCTACATCAATGGAGACAGCAAGCAGTGTATGGGCGAGGTGT 320  
 OY 81 GlyLeuAlaTyrPAlaHisAlaProIleGlyTyrSerLeuSerIleuLeuGlyGlyLeu 100  
 DB 321 GGTCTAGCTGGGCTCAGGCAACCATTTGATATTTCTCTAGTCTAATTTTAGGTCTCG 380  
 OY 101 PhePheAlaLysProMetArgSerIleGlyTyrValThrMetLeuAspProPheIleGln 120  
 DB 381 TTTTTCGCAAACTATGCTTCCAAAGGATATGTACTATGTAGACCATTTCAACAG 440  
 OY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140

DB 441 ATCTATGAAAGCCGATGAGGGGCTGCTCTTCATCCCTGCACTGATGGAGATGTTTC 500  
 OY 141 TPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleValAsp 160  
 DB 501 TGGGCTGAGCAATTTTCTTCGATTTAGGGGCCACCATCAGGTCATCATTTGATG 560  
 OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGly 180  
 DB 561 GTGAACATATGCGTCATTTCTCTGCACTCATTTGCCATTTTAAACCTATGATGGTGG 620  
 OY 181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIle 195  
 DB 621 CTTAAGTGTGGCATATATCATGATGTTTCCAGCTATTTCGATT 665

## RESULT 12

DT207246

LOCUS

DEFINITION

DT207246 843 bp mRNA linear EST 16-AUG-2005  
 JGI\_CAS14763.fwd CAAS Pimphales promelas brain 7-8 month adults,  
 males and females pooled (H) Pimphales promelas cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DOE Joint Genome Institute  
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 Tel: 925 296 5600  
 Fax: 925 296 5710  
 Email: [cdna@jgi-psf.org](mailto:cdna@jgi-psf.org)  
 cDNA Library Preparation: DOE Joint Genome Institute:  
<http://www.jgi.doe.gov>  
 DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>  
 Naming Conventions: EST name is generated by the concatenation of  
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'  
 indicates a forward sequencing read of the insert. It does not  
 necessarily reflect the orientation of the insert.  
 Plate: CAAS 0153 row: F column: 19  
 High quality sequence scop: 754.

## FEATURES

source

1..843

/organism="Pimphales promelas"

/mol\_type="mRNA"

/db\_xref="taxon:90988"

/clone="CAAS14763"

/tissue\_type="brain"

/clone\_lib="CAAS Pimphales promelas brain 7-8 month  
adults, males and females pooled (H)"/note="Vector: pCMVSPORT6; The library was made from dt  
primed cDNA and cloned into Invitrogen vector pCMVSPORT6.  
Poly A RNA were primed with an oligo dt primer (5'-  
GACTAGTCTAGATCGAGCGGCCGCTTTTCTTTTCTTTT -3') ligated  
to a SalI adapter (5'- TCGACCCAGCGCTCG and 5'-  
CGGACGCGGTGG) and digested with NotI. cDNA was size  
selected using 1.1% agarose gel electrophoresis (L  
~0.5-1.2k, M ~1.2-2.5k, H ~2.5k) then ligated into NotI  
and SalI digested pCMVSPORT6 vector. The work was done at  
DOE Joint Genome Institute."

## ORIGIN

## Alignment Scores:

Pred. No.: 5..37e-96 Length: 843

Score: 968.00 Matches: 187  
 Percent Similarity: 88.9% Conservative: 21  
 Best Local Similarity: 79.9% Mismatches: 20  
 Query Match: 32.3% Indels: 6  
 DB: 10 Gaps: 3

US-10-724-806-4 (1-580) x DT207246 (1-843)

QY 1 MetProPheH1ValGluGlyLeuValAlaIleLeuPheTyRleuPheLeu 20  
 DB 144 ATGAACATCATGTGGAGGGGTTGGTGGATTCATCTTCTGATCCTTGA 203  
 QY 21 ValGlyIleTPAlaAlaTPlySThlySaSnSerGlyAsnProGlu-----GluArg 38  
 DB 204 GTGGGATGTGGGCGACATGAGAAAACAAGATTCTGGCGTATGAGGGAATGATCC 263  
 QY 39 SerGluAlaIleIleValGlyArgAspIleGlyLeuLeuValGlyGlyPheThMet 58  
 DB 264 AGTGAGATTATATATGTCGGCGAAGGATATTTGATTTGTCGGTCAATTTACATG 323  
 QY 59 ThrAlaThTPValGlyGlyTyRleAsnGlyThralGluAlaValTyRglyPro 78  
 DB 324 ACTGCAACTGGTGGAGGTGGTATATCAACGCACTGCAGAGATGTTATCTTCCC 383  
 QY 79 GlyCySGlyLeuAlaTPAlaH1saIaProIleGlyTyRSerLeuSerLeuIleGly 98  
 DB 384 GGGTACGGCTTGGCATGGGCGACAGGCTCCCTTTGGATAGCACTGACCTTGTGGGT 443  
 QY 99 GlyLeuPhePheAlaIysProMetArgSerIysGlyTyRValThMetLeuSppPhe 118  
 DB 444 GGCCTTTTCTTGGCTAAACCCATGCGCGGCGGTACGTCACATGCTTGACCCCTTC 503  
 QY 119 LysGlnIleTyRglyArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138  
 DB 504 CAGCAACTGTATGAGAAAAGATGGGCGGTGCTGCTATTCATCCCTCTCAATGGGGGAA 563  
 QY 139 MetPheTPAlaAlaIaIlePheSerAlaLeuGlyAlaThRleSerValIleIleAsp 158  
 DB 564 ATGTTCTGTGTCAGGGCGCATCTTATCTGCGCTTGGGCGACACTAGTGTGATCTTGA 623  
 QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyRleuVal 178  
 DB 624 ATCGACATCAATATGCTGTGATTTATTTGCGCTTATTTCCATTTTCAACGCGGT 683  
 QY 179 GlyGlyLeuTyRSerValaIaTyRThraSppValaIleGlyLeuPheCysIlePheIleGly 198  
 DB 684 GGAGGACTTACTCACTGCTTCACTGATGATGTTGTCAGCTCTTCTGCAATTTTAAAGA 743  
 QY 199 LeuTPAlaSerValProPheAlaLeuSerHisProAlaValThraSppIleGlyPheThr 218  
 DB 744 CTGTGGTGAAGGCTCCCTTTTGCACTCTCAACCCAGCTGTGTCTGATCGAGTGA 803  
 QY 219 AlaVal-----HisAlaIysTyRInSerProTPleu 229  
 DB 804 GCAGTGAACAGCTCATGCTAAT---CAGACTGCATGCTC 842  
 RESULT 13  
 LOCUS DT197042 781 bp mRNA linear EST 16-AUG-2005  
 DEFINITION JGI CAAS5315.fwd CAAS pimephales promelas brain 7-8 month adults,  
 males and females pooled (H) Pimephales promelas cDNA clone  
 CAAS5315 5', mRNA sequence.  
 DT197042  
 DT197042.1 GI:73457093  
 KEYWORDS EST.  
 SOURCE Pimephales promelas  
 ORGANISM Pimephales promelas  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
 Cypriniformes; Cyprinidae; Pimephales.  
 1 (bases 1 to 781)  
 REFERENCE Richardson,P., Lucas,S., Rokhsar,D., Deter,J.C., Ng,D.C.,  
 Brokstein,P. and Lindquist,E.A.

# TITLE JOURNAL COMMENT

DOE Joint Genome Institute Pimephales promelas EST project  
 Unpublished (2005)  
 Other ESTs: JGI CAAS5315.rev  
 Contact: Lindquist,E.A., Richardson,P.  
 DOE Joint Genome Institute  
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 Tel: 925 296 5600  
 Fax: 925 296 5710  
 Email: cdna@jgi-psf.org  
 CDNA Library Preparation: DOE Joint Genome Institute:  
 http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  
 Naming Conventions: EST name is generated by the concatenation of  
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'  
 indicates a forward sequencing read of the insert. It does not  
 necessarily reflect the orientation of the insert.  
 plate: CAAS 0053 row: F column: 10  
 High quality sequence stop: 761.  
 Location/Qualifiers

## FEATURES source

1..781  
 /organism="Pimephales promelas"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:90988"  
 /clone="CAAS5315"  
 /tissue\_type="brain"  
 /clone\_lib="CAAS Pimephales promelas brain 7-8 month  
 adults, males and females pooled (H)"  
 /note="Vector: pCMVSPORT6; The library was made from dt  
 primed cDNA and cloned into Invitrogen vector pCMVSPORT6.  
 Poly A RNA were primed with an oligo dt primer (5'-  
 GACTGTTCTTACATGCGCGCGCCGCTTTTCTTTTCTTTT-3') ligated  
 to a SalI adapter (5'-TCGACCAAGGCTCCG and 5'-  
 CGGACGCGTGGG) and digested with NotI. cDNA was size  
 selected using 1.1% agarose gel electrophoresis (L  
 0.5-1.2k, M -1.2-2.5k, H -2.5k) then ligated into NotI  
 and SalI digested pCMVSPORT6 vector. The work was done at  
 DOE Joint Genome Institute."

## ORIGIN

### Alignment Scores:

Pred. No.: 1,32e-95 Length: 781  
 Score: 964.00 Matches: 186  
 Percent Similarity: 88.8% Conservative: 21  
 Best Local Similarity: 79.8% Mismatches: 20  
 Query Match: 32.2% Indels: 6  
 DB: 10 Gaps: 3

US-10-724-806-4 (1-580) x DT197042 (1-781)

QY 1 MetProPheH1ValGluGlyLeuValAlaIleLeuPheTyRleuPheLeu 20  
 DB 84 ATGAACATCATGTGGAGGGGTTGGTGGATTCATCTTCTGATCCTTGA 143  
 QY 21 ValGlyIleTPAlaAlaTPlySThlySaSnSerGlyAsnProGlu-----GluArg 38  
 DB 144 GTGGGATGTGGGCGACATGAGAAAACAAGATTCTGGCGTATGAGGGAATGATCC 203  
 QY 39 SerGluAlaIleIleValGlyArgAspIleGlyLeuLeuValGlyGlyPheThMet 58  
 DB 204 AGTGAGATTATATATGTCGGCGAAGGATATTTGATTTGTCGGTCAATTTACATG 263  
 QY 59 ThrAlaThTPValGlyGlyTyRleAsnGlyThralGluAlaValTyRglyPro 78  
 DB 264 ACTGCAACTGGTGGAGGTGGTATATCAACGCACTGCAGAGATGTTATCTTCCC 323  
 QY 79 GlyCySGlyLeuAlaTPAlaH1saIaProIleGlyTyRSerLeuSerLeuIleGly 98  
 DB 324 GGGTACGGCTTGGCATGGGCGACAGGCTCCCTTTGGATAGCACTGACCTTGTGGGT 383  
 QY 99 GlyLeuPhePheAlaIysProMetArgSerIysGlyTyRValThMetLeuSppPhe 118  
 DB 384 GGCCTTTTCTTGGCTAAACCCATGCGCGGGTACGTCACATGCTTGAACCCCTTC 443



OY 119 LysGlnIleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138  
Db 444 CAGCAACTGTATGAGAAAAAGATGGCGGTCTGCTATTCATCCCTGCTTAATGGGGGAA 503  
OY 139 MetPheTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158  
Db 504 ATGTTCTGGTCAGGGGCCATCTTAATCTGCGCTTGGGGCCACACTANGTGTGATCATTTAC 563  
OY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178  
Db 564 ATGCACATCAATATGCTGTGATTAATTCGGCTTATTTGACATTTTCTACACGGCTGTT 623  
OY 179 GlyGlyLeuTyrSerValAlaAlaTyrThrAspValValGlnLeuPheCysIlePheIleGly 198  
Db 624 GGAGACTCTACTCAGCGCCCTACACATGATGTTGCCACTCTTCTGCATTTTATTTAGGA 683  
OY 199 LeuTPleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThr 218  
Db 684 CTGTGGGTGAGCGTCCCTTTTGCACTCTCAAAACCGACTGTGTGACATCGGAGTGACA 743  
OY 219 AlaValIleHisAlaLysTyrGlnSerProTyr 228  
Db 744 GCAGTGAAGCAGCTCCATCTAAT---CAGACTGCATGG 779

RESULT 14  
DT972151 1114 bp mRNA linear EST 21-SEP-2005  
LOCUS CLJ138-H03 5', mRNA sequence.  
DEFINITION DT972151  
ACCESSION DT972151  
VERSION DT972151.1 GI:76044969  
KEYWORDS EST.

SOURCE Gastrosteus aculeatus (three spined stickleback)  
ORGANISM Gastrosteus aculeatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorphi; Acanthopterygii; Percomorpha; Gastrosteiformes;  
Gasterosteidae; Gastrosteus.

REFERENCE 1 (bases 1 to 1114)  
AUTHORS Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,  
Schmutz,J. and Myers,R.M.  
TITLE Expressed sequence tags from Gastrosteus aculeatus  
JOURNAL Unpublished (2003)  
COMMENT Contact: Grimwood, Jane  
Stanford Human Genome Center  
Stanford University School of Medicine  
975 S California Ave, Palo Alto, CA 94304, USA  
Tel: 650 320 5917  
Fax: 650 320 5801  
Email: jane@shgc.stanford.edu  
Plate: 138

FEATURES  
source High quality sequence stop: 758.  
Location/Qualifiers  
1..1114  
/organism="Gastrosteus aculeatus"  
/mol\_type="mRNA"  
/strain="Blitrufojordu marine sticklebacks, Iceland"  
/db\_xref="taxon:69293"  
/clone="CLJ138-H03"  
/sex="mixed male and female"  
/tissue\_type="whole larva"  
/dev\_stage="21 day old larvae collected at Swarnup Stage 30  
(J. Embryol. Exp. Morphol 6: 373-383,1958)"  
/lab\_host="DH10B (T1 phage resistant)"  
/note="Vector: Express 1; Total and poly A+ RNA was  
isolated from the indicated stickleback tissue, and a cDNA  
library was constructed in the Express 1 plasmid vector by  
Open Biosystems. First strand cDNA synthesis was primed  
with an 54 bp linker primer containing an oligodT sequence  
preceded by a synthetic NotI site (first strand primer:  
5'-GACTAGTTCTAGATCGGAGCGCGCC(TT25-3')-. Following  
second strand synthesis, cDNAs were made blunt at the end

corresponding to the original 5 prime end of mRNA, and  
cloned directionally into the NotI and EcoRV sites of  
Express 1. Note that the EcoRV site is typically destroyed  
in the blunt end cloning, leaving a junction of the form  
'xxxAATC' (where is ATC is the second half of the EcoRV  
site, and xxx is derived from the cDNA sequence). A map of  
the Express 1 vector is available at:  
[http://www.openbiosystems.com/cdna\\_library\\_construction\\_fa](http://www.openbiosystems.com/cdna_library_construction_fa)  
q.php#8 The primary library was transformed and amplified  
in DH10B (T1 phage resistant) bacteria. Clones available  
from Open Biosystems:  
<http://www.openbiosystems.com/stickleback>

## ORIGIN

Alignment Scores:

Pred. No.:	3,82e-94	Length:	1114
Score:	953.00	Matches:	183
Percent Similarity:	80.2%	Conservative:	27
Best Local Similarity:	69.8%	Mismatches:	50
Query Match:	31.8%	Indels:	2
DB:	10	Gaps:	1

US-10-724-806-4 (1-580) x DT972151 (1-1114)

OY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
Db 278 ATGACCATCCACATAGAGGGCTGTGCTATCGGCACTTCTATCTGATCATCTGTGC 337  
OY 21 ValGlyIleTPAlaAlaIleTyrThrIleSerValIleIleAspGlu-----GluArg 38  
Db 338 GTGGGCATCTGGCGCGCATGAAAAACAACACTCCGGGAGCGCGGAGCCGACCGACCGC 397  
OY 39 SerGlnAlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMet 58  
Db 398 AGCGAAACCATCATGCTGCTGGTGGAGGACATGGATTATTTGTTGGGATTTACATG 457  
OY 59 ThrAlaThrTPValGlyGlyTyrIleLeuGlyThrAlaGluAlaValTyrGlyPro 78  
Db 458 ACAGGACGTGGGTGGAGGAGATATTAATGATACAGCTGAGTATGTATCTGCCT 517  
OY 79 GlyCysGlyLeuAlaTPAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGly 98  
Db 518 GCGTACGCTTGGCTTGGGCTCAAGCTCCCTTGGATATGACTGACTGCTTGTGTGGG 577  
OY 99 GlyLeuPhePheAlaLysProMetArgSerIleGlyTyrValThrMetLeuAspProPhe 118  
Db 578 GGTATATTTTGGCAAGCCCATCGCTCAAGAGTTACGTACCATGTGAGACCGCTTC 637  
OY 119 LysGlnIleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138  
Db 638 CAGCAGATCTTACGGGAAACGATGGGTGCTCTCTTCATACCTGCACATGAGGAGAG 697  
OY 139 MetPheTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158  
Db 698 ATCTTCTGGTCGGCGCCATCTTAATCCGCCCTCGGTGCTACTGATGTATCATCTGGAGC 757  
OY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuVal 178  
Db 758 ATNCACATTAAGATGCTCGGTGTTATCTCAGCGCTCATTTTACATTTTACACCTTGGTT 817  
OY 179 GlyGlyLeuTyrSerValAlaAlaTyrThrAspValValGlnLeuPheCysIlePheIleGly 198  
Db 818 GGNACACTTACTCGGTGCTCACACCATATGTCGAGAGCTTTTGCAATCTTCTTGGN 877  
OY 199 LeuTPleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThr 218  
Db 878 CTGTGATCAATACGTCCCGTTGCTGACACACCCCTGGGTTCTGACATCATCATTAAT 937  
OY 219 AlaValHisAlaLysTyrGlnSerProTyrLeuGlyThrIleGlySerValGluValTyr 238  
Db 938 GCAGTGAACAAGTGTACATCGCTTGAAGAAACCAACCTCCCAAGGACCCCGGG 997  
OY 239 ThrTPleuAspAsnPheLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPhe 258



Db 998 GTGGGAGCAATACCTCCGCTGTGATGCTTGAAGAATCCCTGGGAAGGATTTTC 1057  
 QY 259 GlnArg 260  
 Db 1058 CAAAAA 1063

RESULT 15  
 DM614674  
 LOCUS DM614674  
 DEFINITION CLJ282-809, y1d-s SHGC-CLJ2 Gasterosteus aculeatus cDNA clone  
 ACCESSION DM614674  
 VERSION DM614674.1 GI:85217654  
 KEYWORDS EST.  
 SOURCE Gasterosteus aculeatus (three spined stickleback)  
 ORGANISM Gasterosteus aculeatus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.

REFERENCE 1 (bases 1 to 1054)  
 AUTHORS Kingley, D.M., Peichel, C., Knecht, A., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.  
 TITLE Expressed sequence tags from Gasterosteus aculeatus (2004)  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Grimwood, Jane  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 975 S California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 320 5917  
 Fax: 650 320 5801  
 Email: jane@shgc.stanford.edu  
 Plate: 282  
 High quality sequence stop: 875.  
 Location/Qualifiers  
 1. 1054  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="mRNA"  
 /strain="Bittorfjordur marine sticklebacks, Iceland"  
 /db\_xref="taxon:69293"  
 /clone="CLJ282-809"  
 /sex="mixed male and female"  
 /tissue\_type="whole larva"  
 /dev\_stage="21 day old larvae collected at Swarnup Stage 30 (J. Embryol. Exp. Morphol 6: 373-383, 1998)"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1ib="SHGC-CLJ2"  
 /note="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTGATTCATGATCGGAGCGGCCCT(T)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: http://www.openbiosystems.com/cdna\_library\_construction\_fa q.php# The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria."

ORIGIN

Alignment Scores:  
 Pred. No.: 3 38e-87 Length: 1054  
 Score: 890.00 Matches: 168  
 Percent Similarity: 68.8% Conservative: 75  
 Best Local Similarity: 47.6% Mismatches: 106

Query Match: 29.7% Indels: 4  
 DB: 10 Gaps: 3  
 US-10-724-806-4 (1-580) x DM614674 (1-1054)

QY 73 GUAAlaValTYrGlyProGlyCysGlyLeuAlaTPAlaHisAlaProIleGlyTYrSer 92  
 Db 2 GAGGCAACATACACCCGACACTGGCGGGGGTGGGCCCTCATGGCTGTGCAATATGTC 61

QY 93 LeuSerLeuIleLeuGlyLeuPhePheAlaValProMetArgSerLeuGlyTYrVal 112  
 Db 62 CTGACCTCTTCTTAgTGGGTTTTCCTTGCCAGCCTTATGAGAGAACAGATATGTA 121

QY 113 ThrMetLeuAspProPheLeuGlySerValIleTYrGlyValArgMetGlyLeuLeuPheIle 132  
 Db 122 ACAATATGAGACCCCTTCCACAAAGTATGGGAACGTTGTAGAGCGTGGTTCCTT 181

QY 133 ProAlaLeuMetGlyGluMetPheTPAlaAlaIlePheSerAlaLeuGlyAlaThr 152  
 Db 182 CCGGCTCTGTGGCTGATGTCTGTGGTGGGACGACACACTGTCTGACCTGGGTGGGACC 241

QY 153 IleSerValIleLeuAspValAspValIleSerValIleValSerAlaLeuIleAla 172  
 Db 242 ATGAGGTATCTGAGACCTGCTCCTACGCTACTCTCATATCATCTCTGAGCGGTGGC 301

QY 173 IleLeuTYrThrLeuValGlyGlyLeuTYrSerValAlaTYrThrAspValValGlnLeu 192  
 Db 302 ATGATCTACACACTGCTGGGGGGGCTCTACTCTGTGGCCATAGGACGTCATCCAGCTC 361

QY 193 PheCysIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaVal 212  
 Db 362 ATCCATCTTGTGTGCTGCTGGGTGGTGTGCTCTTCTCTGATGACCAACCTCCTCTCT 421

QY 213 ThrAspIleGlyPheThrAlaValHisAlaValTYrGlySerProTrpIleGlyThrIle 232  
 Db 422 ATGACATCTCACTGACGCTGACGCTACACGAGACCTTCAAGCTCTCTGGGTGGACAGTG 481

QY 233 GluSerValGluValTYrThrTrpLeuAspAspPheLeuLeuMetLeuGlyGlyIle 252  
 Db 482 GAGCTGATGAGGCGCGGCAAGTGCTTCGATGATTCATGCTGTGGCTCGGGGGGTTTG 541

QY 253 ProTrpGlnAlaTYrPheGlnArgValLeuSerSerSerSerAlaThrTYrAlaGlnVal 272  
 Db 542 GCTTACAGGCGTTTATCCAAAGATTTATCCGCTCGCTTATCACCCAGGCGCCAGTT 601

QY 273 LeuSerPheLeuAlaIlePheGlyCysLeuValMetAlaLeuProAlaIleCysIleGly 292  
 Db 602 ACCTGTCTGCTCTCCAGGCTTCTGCTGGTGGTGGGATCCCTTCATTTCTGGTGGG 661

QY 293 AlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTYrGlyTYrProAspProLYrThr 312  
 Db 662 GCTGTGGCCGCTCTTACAGACTGGAATCTCAACAGCTATGATTAACCAACCATATGAA 721

QY 313 LysGlnGlnAlaAspMetIleLeuProIleValLeuGlnIleCysProValTYrIle 332  
 Db 722 CGCAACAGGACAGGCTCAATCTCCCATGCGCTCGATGCTTCACACCACTCACTAGTC 781

QY 333 SerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSerIle 352  
 Db 782 TCCATCATCGGATGGAGTGAAGTGAAGTGGCGCATGTGCTCGATGAGTGAAGTGAAGT 841

QY 353 LeuSerAlaSerSerMetPheAlaArgAsnIleTYrGlnLeuSerPheArgGlnAsnAla 372  
 Db 842 CTGTCTCTGTCATCATTTGTTTCTTCAAAATTTATTCACAGAACATCCCTCAGAACAGCA 901

QY 373 SerAspLysGluIleValTrpValMetArgIleThrValLeuValPheGlyAlaSerAla 392  
 Db 902 TCCGAACGTGAGATGATGGGTGATCCGATCCCTCGTGTGATGATGGGTGTGCTGG 961

QY 393 ThrAlaMetAlaLeuLeuThr---LysThrValTYrGlyLeuTrpTYrIleSerSerAsp 411  
 Db 962 CACGCCCTTACCTTTGAGACACAGCGTTCTGTTCTC-----TGGCTGTGGGCGTGGAC 1015

Oy 412 LeuValTyrIleIleIlePheProGlnLeuLeuCySval 424  
::: ||||| :::  
Db 1016 ATGCTACACATCATGTT---CCTCAGCGGTGCTGGGTC 1051

Search completed: July 10, 2006, 16:12:01  
Job time : 7697 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: July 10, 2006, 14:15:32 ; Search time 1898 Seconds  
(without alignments)  
5632.372 Million cell updates/sec

Title: US-10-724-806-4  
Perfect score: 2993  
Sequence: 1 MPFHEGLVAILFLIFL.....EALLDVDSSEGGSTEDNLQ 580

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues  
Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_p2n.model -DEV=xlh  
-O=/abss/ABSSWEB\_spool/US10724806/runat\_10072006\_140255\_7855/app\_query.fasta.1  
-DB=Published\_Applications\_NA\_Main -OFMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext  
-HEBESIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=dbse07  
-USER=US10724806@CEN\_1\_1\_1675@runat\_10072006\_140255\_7855 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEE\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_Main:

1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11E\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2993	100.0	1743	3	US-09-911-077A-3 Sequence 3, Appl1

2	2993	100.0	1743	3	US-09-911-077A-23	Sequence 23, Appl1
3	2993	100.0	1743	9 <th>US-10-724-806-3</th> <th>Sequence 3, Appl1</th>	US-10-724-806-3	Sequence 3, Appl1
4	2993	100.0	1743	9 <th>US-10-724-806-23</th> <th>Sequence 23, Appl1</th>	US-10-724-806-23	Sequence 23, Appl1
5	2993	100.0	1743	3 <th>US-09-911-077A-5</th> <th>Sequence 5, Appl1</th>	US-09-911-077A-5	Sequence 5, Appl1
6	2993	98.8	4804	9 <th>US-10-724-806-5</th> <th>Sequence 5, Appl1</th>	US-10-724-806-5	Sequence 5, Appl1
7	2795	93.4	1743	3 <th>US-09-911-077A-1</th> <th>Sequence 1, Appl1</th>	US-09-911-077A-1	Sequence 1, Appl1
8	2795	93.4	1743	9 <th>US-10-724-806-1</th> <th>Sequence 1, Appl1</th>	US-10-724-806-1	Sequence 1, Appl1
9	2795	93.4	1813	3 <th>US-09-911-077A-9</th> <th>Sequence 9, Appl1</th>	US-09-911-077A-9	Sequence 9, Appl1
10	2795	93.4	1813	9 <th>US-10-724-806-9</th> <th>Sequence 9, Appl1</th>	US-10-724-806-9	Sequence 9, Appl1
11	1491.5	49.8	1833	8 <th>US-10-241-784-1</th> <th>Sequence 1, Appl1</th>	US-10-241-784-1	Sequence 1, Appl1
12	1444.5	48.3	1885	3 <th>US-09-911-077A-7</th> <th>Sequence 7, Appl1</th>	US-09-911-077A-7	Sequence 7, Appl1
13	1444.5	48.3	1985	9 <th>US-10-724-806-7</th> <th>Sequence 7, Appl1</th>	US-10-724-806-7	Sequence 7, Appl1
14	1413	47.2	4223	13 <th>US-11-097-143-41593</th> <th>Sequence 41593, A</th>	US-11-097-143-41593	Sequence 41593, A
15	1405	46.9	1729	13 <th>US-11-097-143-41594</th> <th>Sequence 41594, A</th>	US-11-097-143-41594	Sequence 41594, A
16	1099.5	36.7	1461	3 <th>US-09-974-300-501</th> <th>Sequence 501, App</th>	US-09-974-300-501	Sequence 501, App
17	1096	36.6	26323	15 <th>US-11-005-030-1</th> <th>Sequence 1, Appl1</th>	US-11-005-030-1	Sequence 1, Appl1
18	1096	36.6	119040	3 <th>US-09-911-077A-19</th> <th>Sequence 19, Appl1</th>	US-09-911-077A-19	Sequence 19, Appl1
19	1096	36.6	119040	9 <th>US-10-724-806-19</th> <th>Sequence 19, Appl1</th>	US-10-724-806-19	Sequence 19, Appl1
20	1096	36.6	142299	3 <th>US-09-911-077A-14</th> <th>Sequence 14, Appl1</th>	US-09-911-077A-14	Sequence 14, Appl1
21	1096	36.6	142299	9 <th>US-10-724-806-14</th> <th>Sequence 14, Appl1</th>	US-10-724-806-14	Sequence 14, Appl1
22	318	10.6	2475	8 <th>US-10-302-172-932</th> <th>Sequence 932, App</th>	US-10-302-172-932	Sequence 932, App
23	314	10.5	9025608	7 <th>US-10-156-761-1</th> <th>Sequence 1, Appl1</th>	US-10-156-761-1	Sequence 1, Appl1
24	313	10.5	2105	10 <th>US-10-055-877-116</th> <th>Sequence 116, App</th>	US-10-055-877-116	Sequence 116, App
25	313	10.5	2105	12 <th>US-10-455-772-911</th> <th>Sequence 911, App</th>	US-10-455-772-911	Sequence 911, App
26	313	10.5	2105	12 <th>US-10-455-772-917</th> <th>Sequence 917, App</th>	US-10-455-772-917	Sequence 917, App
27	313	10.5	2200	10 <th>US-10-874-706-56</th> <th>Sequence 56, App</th>	US-10-874-706-56	Sequence 56, App
28	313	10.5	2384	9 <th>US-10-757-262-93</th> <th>Sequence 93, Appl1</th>	US-10-757-262-93	Sequence 93, Appl1
29	313	10.5	2384	9 <th>US-10-768-158-45</th> <th>Sequence 45, Appl1</th>	US-10-768-158-45	Sequence 45, Appl1
30	312	10.4	2028	3 <th>US-09-733-630-1</th> <th>Sequence 1, Appl1</th>	US-09-733-630-1	Sequence 1, Appl1
31	312	10.4	2456	3 <th>US-09-733-630-3</th> <th>Sequence 3, Appl1</th>	US-09-733-630-3	Sequence 3, Appl1
32	310	10.4	2028	3 <th>US-09-928-530-3</th> <th>Sequence 3, Appl1</th>	US-09-928-530-3	Sequence 3, Appl1
33	310	10.4	2028	6 <th>US-10-162-012-28</th> <th>Sequence 28, Appl1</th>	US-10-162-012-28	Sequence 28, Appl1
34	310	10.4	2028	6 <th>US-10-162-102-28</th> <th>Sequence 28, Appl1</th>	US-10-162-102-28	Sequence 28, Appl1
35	310	10.4	2326	3 <th>US-09-928-530-1</th> <th>Sequence 1, Appl1</th>	US-09-928-530-1	Sequence 1, Appl1
36	310	10.4	2326	6 <th>US-10-162-012-26</th> <th>Sequence 26, Appl1</th>	US-10-162-012-26	Sequence 26, Appl1
37	310	10.4	2326	7 <th>US-10-162-102-26</th> <th>Sequence 26, Appl1</th>	US-10-162-102-26	Sequence 26, Appl1
38	308	10.3	1458	7 <th>US-10-156-761-5268</th> <th>Sequence 5268, App</th>	US-10-156-761-5268	Sequence 5268, App
39	307	10.3	2043	9 <th>US-10-757-262-1</th> <th>Sequence 1, Appl1</th>	US-10-757-262-1	Sequence 1, Appl1
40	307	10.3	2107	8 <th>US-10-302-172-234</th> <th>Sequence 234, App</th>	US-10-302-172-234	Sequence 234, App
41	305	10.2	3493	16 <th>US-11-136-537-2220</th> <th>Sequence 2220, App</th>	US-11-136-537-2220	Sequence 2220, App
42	301	10.1	142299	3 <th>US-09-911-077A-14</th> <th>Sequence 14, Appl1</th>	US-09-911-077A-14	Sequence 14, Appl1
43	301	10.1	142299	9 <th>US-10-724-806-14</th> <th>Sequence 14, Appl1</th>	US-10-724-806-14	Sequence 14, Appl1
44	300	10.0	1974	12 <th>US-10-535-066-53</th> <th>Sequence 53, Appl1</th>	US-10-535-066-53	Sequence 53, Appl1
45	298	10.0	1413	8 <th>US-10-282-122A-16541</th> <th>Sequence 16541, A</th>	US-10-282-122A-16541	Sequence 16541, A

## ALIGNMENTS

RESULT 1

US-09-911-077A-3  
; Sequence 3, Application US/09911077A  
; Publication No. US20030114339A1

GENERAL INFORMATION:

APPLICANT: BLAKELEY, RANDY D.

APPLICANT: APPASUNDARAM, SUBRAMANIAM

APPLICANT: FERGUSON, SHAWN

TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA

FILE REFERENCE: VBLT:00808

CURRENT APPLICATION NUMBER: US/09/911,077A

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 1743

TYPE: DNA

ORGANISM: Mus musculus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1743)

US-09-911-077A-3

Alignment Scores:

Pred. No.:

2.33e-300

Length:

1743

Score: 2993.00 Matches: 580  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-3 (1-1743)

```
QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheThrLeuLeuIlePheLeu 20
DB 1 ATGCTTTTCCAGTGGAGGAGCTGGTAGCTATTATCTCTTCTTACCTCTTATTTCTCG 60
QY 21 ValGlyIleTPAlaAlaTrpIysThrIysAsnSerGlyAsnProGluIuArgSerGlu 40
DB 61 GTTGAATATGGGCTGCATGGAACCAAAAACAGGGGCAACCCAGAAAGCGCAGTGAA 120
QY 41 AlaIleIleValGlyValArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB 121 GCCATCATAGTGGGGGCGGTGACATGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
QY 61 ThrTPValGlyGlyValArgIleAsnGlyThrAlaGluAlaValIysGlyProGlyCys 80
DB 181 ACCTGGGTTGGAGGAGGCTACATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 81 GlyLeuAlaTPAlaHisAlaProIleGlyIysSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 241 GGTCTAGGCTGGGCTCATGACACCATGGATGTCTGTGAGTCATATTTAGGTGGTCTG 300
QY 101 PhePheAlaIysProMetArgSerIysGlyValIleThrMetLeuAspProPheIysGln 120
DB 301 TTTTGGGAAACCTATAGCGTTCCAGAGGATATGACTATATGAGACCATTCAGAAACG 360
QY 121 IleTyrGlyIysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB 361 ATCTATGGAAGCGCATGGGTGGGCTCTTCATCCCGACGCTAGTGGAGAGAGTTC 420
QY 141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGCAGCAATTTCTCTGCATTAAGGGGCCACCATGAGCGTATCATATGATGGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB 481 GTGAACATATCGGTATGTCTCTGCATCTTGCATTTCTTATATACCTTAGGGGTGG 540
QY 181 LeuTyrSerValAlaIleTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 541 CTCTACTGTGGCATATATCTGATGTGTCCAGCTATTTCTGCAATTTTATAGACGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 601 ATCATGTGCTCTTTGGCCCTGTGCACATCTGTGACATCCGACATGGAGTTCACAGCTGG 660
QY 221 HisAlaIysTyrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrThrTrp 240
DB 661 CATGCTTAATACCAAGATCTCTGGCTGGAGAACCATTTGAATCGTTGAAGTCTACACTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaIleTyrPheGlnArg 260
DB 721 CTGTATATATTTCTGTATTATTATGCTGGGTGGAATCCCATGGCAAGCTTATCCAGAGG 780
QY 261 ValLeuSerSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 781 GTCTCTCTCTCATCTTCAGCCACCTATGCTGAGGATCTGTCTTCTGTGACACTTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaIleSerThrAspTrp 300
DB 841 TGCCCTGGTGGTGGCTCTACCCGCCCATATGATGAGAGCTATTGGAGCTTCCACAGACTGG 900
QY 301 AsnGlnThrAlaIleTyrGlyTyrProAspProIysThrIysGlnGluAlaAspMetIleLeu 320
DB 901 AACCAAGACTGGCTTACGGGTATCCAGATCCCAAGACTAAGAGAGAGCAATGATTTCTC 960
QY 321 ProIleValLeuGlnIleTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
```

```
DB 961 CCGATCGTTTGCAGTACCTCTGCGCTGTGATATCTCTTCTTGGGCTGGTGGCTGT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TTAGCTGCTGTATGTTCTCTACGACTGACTGTCCATCTCTGTGGCGAGTCTTATTTTGGCT 1080
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIysGluIleValTrpVal 380
DB 1081 CGGAATATCTACAGACTTCTCTTCCAGCAAAATGATCAGACAGAAAGAAATGTGGGGTTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIys 400
DB 1141 ATGAGATCATCTGTCTGTGTGGACATCTGCACAGCCATGCGCTTGTGTGACGAAG 1200
QY 401 ThrValTyrGlyLeuTrpTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1201 ACTGTATGGGCTTGGTGTGACTGAGCTCTGACCTTGTCTTACATATATCTTCCACAG 1260
QY 421 LeuLeuCysValLeuPheIleIysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1261 CTGCTCTGTGATCTCTTCATCATAGAGAACCAACTTATGGGGCAGTGTGTATATTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460
DB 1321 TTTGACATATCTCTGAGATTAATGAGAGAGAGCCATATCTATACTTACAGCCCTTAATC 1380
QY 461 PheTyrProGlyTyrTyrSerAspIysAsnGlyIleTyrAsnGlnArgPheProPheIys 480
DB 1381 TTCTACCTGTGTATTTACTCTGCACAGAAATGTATATCAATCAGAGGTTCCCATTTAAA 1440
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIysTyr 500
DB 1441 ACTCTCTCATAGTTACTCATCTTTTACCAAACTTGTGTCTTACTGACCAAGTAT 1500
QY 501 LeuPheGlnSerGlyThrLeuProProIysLeuAspValPheAspAlaValAlaIleArg 520
DB 1501 CTATTTGAAAGTGAACCTTGCTCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
QY 521 HisSerGlnGluAsnMetAspIysThrIleLeuValArgAsnGluAsnIleIysLeuAsn 540
DB 1561 CACAGTAGAGACATGTGACAGACCACTTCTAGTCAGAAATGAAATATCAATTAATTAAT 1620
QY 541 GluLeuAlaProValIysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIys 560
DB 1621 GAACCTGACCTGTGAACCTCGGAGAGCTTAACCTCACTTCACTTCACTTCACTTCACT 1680
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB 1681 GAGGCTCTCTGTGATGTGATTCAGTCCGAGAGGGGTCTGGAGACTGAAGATATATTACAA 1740

RESULT 2
US-09-911-077A-23
; Sequence 23, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBTL:008US
; CURRENT APPLICATION NUMBER: US/09/911.077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-23
```

## Alignment Scores:

Pred. No.:	2,33e-300	Length:	1743
Score:	2993.00	Matches:	580
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-724-806-4 (1-580) x US-09-911-077A-23 (1-1743)

```

QY      1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB      1 ATGGCTTTCATGTGGAGAGACTGTAGCTATTATCTCTTACCTCTTATATTCTG 60
QY      21 ValGlyIleTyrAlaIleTyrPheTyrLeuSerGlyAsnProGluGlnAgsSerGlu 40
DB      61 GTTGGAAATGGGCTGCATGAAACCAAAACAGGGCAACCCAGAGAGCGCATGTGA 120
QY      41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB      121 GCCATATAGTGGGGGCGGCGCTGACATTTGTTGTTGGTGGTTTACCATGACAGCC 180
QY      61 ThrTyrValGlyGlyTyrIleAsnGlyThrAlaGlnAlaValTyrGlyProGlyCys 80
DB      181 ACTGGGTTGGAGAGGCTACATCATGGGACAGAGAGAGAGAGTATGGGCGAGGTTGT 240
QY      81 GlyLeuAlaTyrAlaIleHisAlaProIleGlyTyrSerLeuSerIleLeuGlyGlyLeu 100
DB      241 GGCTTACCTTGGGCTCATGACCCCATGTGATATCTGATCTGATTTTAAAGTGTCTG 300
QY      101 PhePheAlaIlePheMetArgSerIleGlyTyrValThrMetLeuAspProPheLeuGln 120
DB      301 TTTTTCGCAAACTTATGCGTTCCAGGATATGTACATATGTAAACCATTTCAACAG 360
QY      121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
DB      361 ATCTATGAAAGGAGGCTAGGTGGGCTGCTCTTCATCCCTGACATGATGGAGAGATGTT 420
QY      141 ThrAlaIleAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB      421 TGGGCTGAGCAATTTCTCTGCAATTAGGGCCACATCAGCGTGAATGATGGAGAT 480
QY      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB      481 GTGACATATTCGGTCTATTTGCTCTGCACTCATTTGCACTTTTAACTTAACTTAACT 540
QY      181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
DB      541 CTCTACTCTGTGGCATATATCTGATGTGTCACACTATTTGCACTTTTAAAGACTGTGG 600
QY      201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB      601 ATCAGAGTCCCTTTTCCCTGTCACATCTGACAGTACCGACATTCGATTTCAAGCTGTG 660
QY      221 HisAlaIleTyrGlnSerProTyrLeuGlyThrIleGlySerValGlnValTyrThrTyr 240
DB      661 CATGCTAAATCCAGAGTCCCTGCTGGGAGAACCATTTGAATCAAGCTTCAAGCTGG 720
QY      241 LeuAspAsnPhelLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260
DB      721 CTGTATTAATTTCTGTATTAATGATGTGGGGAATCCCATGGAGACCTTACTTCCAGAG 780
QY      261 ValLeuSerSerSerSerAlaTyrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB      781 GTCTCTCTTCACTCCAGCCACCTAAGCTCAGGTAAGCTCTTCTTCCGAGGCTTTTGGG 840
QY      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
DB      841 TGGCTGTGATGCTTACCCGCGCATATGATGAGAGCTATTGAGGTTTCAAGAGCTGG 900
QY      301 AsnGlnThrAlaTyrGlyTyrProAspProGlyThrIleGlyGlnAlaIlePheIleLeu 320

```

```

DB      901 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTTAAGAGAGAGACATGATTCTC 960
QY      321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB      961 CCGATGTTCTGAGTACCTTCTGCGCTGTGTACATCTCTTCTTGGGCTGGTGTCTGTT 1020
QY      341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB      1021 TCAGTGTGCTGATGCTTCACTGATGCTGCTGATCTGCTGATCTGCTGATGCTGCT 1080
QY      361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal 380
DB      1081 CGGAATATCTACAGCTTCTTCCACAAATATGATCAGACAGAGAAATGTGTGGGCTC 1140
QY      381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB      1141 ATGAGGATCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
QY      401 ThrValTyrGlyLeuTyrPheTyrLeuSerSerAspLeuValTyrIleIlePheProGln 420
DB      1201 ACTGTATAGGCTGTGATACCTGAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 1260
QY      421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB      1261 CTGCTCTGTGATCTTTCATCAAAAGAGAACCACTTATGGGCGAGTGTGCTGTTATTT 1320
QY      441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460
DB      1321 TTGGACTATCTCTGAAATTAAGTGAAGAGAGCCATATCTATCTTGGAGCCCTTAATC 1380
QY      461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
DB      1381 TTCTACCTGTTATTAATCTGACAGAAATGATATTAATCAATCAGAGGTTCCCATTTAA 1440
QY      481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
DB      1441 ACTCTCTCATGTTTACCTTCATCTTCTTACCAATTTGTTGTTCTTATCTAGCAAGTAT 1500
QY      501 LeuPheGlnSerGlyTyrThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
DB      1501 CTATTGAAAGTGAACCTTGGCTCCAAATTAAGATTAATGATGCTGTGTGGCAGAG 1560
QY      521 HisSerGluLysAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
DB      1561 CACAGTGAAGAGACATGACAAACCATTTAGTACAGAAATGAATATCAAAATTAAT 1620
QY      541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB      1621 GAACCTGACACTGTGAACCTCGGAGAGCCCTAACCTTCAGTTCAACTTCAACAAATAAG 1680
QY      561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB      1681 GAGGCCCTCTTGATGTTGATTTCCAGTCCGAGGGGCTGGGACTGAAGATTAATTAACA 1740

```

## RESULT 3

US-10-724-806-3

Sequence 3, Application US/10724806

Publication No. US20040248838A1

GENERAL INFORMATION:

APPLICANT: BLAKELY, RANDY D.

APPLICANT: APPASUNDARAM, SUBRAMANIAM

TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA

CURRENT APPLICATION NUMBER: US/10/724,806

PRIOR FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: US/09/911,077A

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3

LENGTH: 1743

TYPE: DNA

```
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1743)
US-10-724-806-3

Alignment Scores:
  Pred. No.:      2,336-300      Length:      1,743
  Score:          2993.00        Matches:      580
  Percent Similarity: 100.0%      Conservative: 0
  Best Local Similarity: 100.0%    Mismatches: 0
  Query Match:      100.0%        Indels:      0
  DB:                9           Gaps:          0

US-10-724-806-4 (1-580) x US-10-724-806-3 (1-1743)

QY      1 MetProPheHisValGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB      1 ATGCCTTCCATGTCGGAAGACTGGTAGCTATTATCTCTTACTCTCTTATATTCTG 60
QY      21 ValGlyIleTPrAlaAlaTPrIysThrIysAsnSerGlyAsnProGluGluArgSerGlu 40
DB      61 GTTGGAATATGGGCTGCATGCAAAACCAAAACAGCGGCAACCCAGAGAGCGCAGTGA 120
QY      41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
DB      121 GCCATCATATGCGGGGCGGTGACATGGTTGGTTGGTTGGTTTAAACCATGACAGCC 180
QY      61 ThrTPrValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB      181 ACGTGGTGGAGAGAGGCTACATCAATGGAGACAGCAAGACGTATATGGCCAGGTTGT 240
QY      81 GlyLeuAlaTPrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeu 100
DB      241 GGTACGCTGGGCTCATGACACCCATGGATATCTCTGAGCTAATTTAGTGGTCTG 300
QY      101 PhePheAlaIysProMetArgSerIysGlyTyrValThrMetLeuAspProPheIysGlu 120
DB      301 TTTTGGGAAACCTATGGGTTCCAAAGGATATGACATATGTATAGACCATTCAAACG 360
QY      121 IleTyrGlyLysArgMetGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB      361 ATCTATGGAAGCGGATGGGTGGGCTGCTCTTCACTCCGCACTGATGGGAGAGATGTC 420
QY      141 TPrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB      421 TGGGCTGCAGCAATTTCTCTGCAATTAGGGGCCACCATCAGCGTATCATGATGTGAT 480
QY      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB      481 GTGAACATATGCGTCAATGCTCTGCACTGCACTGCAATTCCTTATTAACCTAGTGAGTGG 540
QY      181 LeuTyrSerValAlaIleTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTPr 200
DB      541 CTCCTACTGTGGCATATACGATGTTGCCAGCATTTCTGCACTTTTATAGAGACTGAG 600
QY      201 IleSerValProPheAlaLeuSerHisProAlaValIleThrAspIleGlyPheThrAlaVal 220
DB      601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTCGACATCCGACGATCGAATTCAGATTCAG 660
QY      221 HisAlaLysTyrGlnSerProTPrLeuGlyThrIleGluSerValGluValTyrThrTPr 240
DB      661 CATGCTTAATATACAGAGTCCCTGGGCAACCATTAATCACTGTAAGTCTACACCTGG 720
QY      241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTPrGluAlaIleTyrPheGlnArg 260
DB      721 CTGTATATATTTCTCTATATGATGCTGGGTGGAATCCCATGCAAGCCACTCTCAAGAG 780
QY      261 ValLeuSerSerSerSerAlaThrTyrAlaGluValLeuSerPheLeuAlaIlePheGly 280
DB      781 GTCCCTCTTATCTCTCAAGCCCATATGCTCAGGATACGTCTCTCTGCACTTTTGGG 840
QY      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTPr 300
```

```
|||||
DB      841 TGCCGTGATGAGGCTCTACCCGCCATATGATAGAGGCTATTTGAGCTTCCACAGACTGG 900
QY      301 AsnGlnThrAlaTyrGlyTyrProAspProIysThrIysGluGluAlaAspMetIleLeu 320
DB      901 AACAGACTGCTTACGGGTATCCAGATCCCAAGACTAAGAGAGAGACAGATGATTTCTC 960
QY      321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB      961 CCGATCGTTTGCAGTATCTCTGCGCCCTGTGAACATCTCTTTTGGGCTTGCTGCTGT 1020
QY      341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB      1021 TCAGCTGCTGCATATCTCTCAGCTGACCTGCTTCACCTGCTGCGGCGAGTTCTATGTTGCT 1080
QY      361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTPrVal 380
DB      1081 CGGAATATCTACAGCTTTCTCTTCCAGACAAATGATATGACACAGCAATGCTTGCTGCAAG 1140
QY      1141 ATGAGATCAGTGTGCTTGTGTCGAGCATCTGCACAGCCATGCTTGTGCTGCAAG 1200
DB      1201 ThrValTyrGlyLeuTPrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
QY      421 LeuLeuCysValLeuPheIleIleGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB      1261 CTGCTCTGTATCTTCTTATCAAGAGAACAACTTATGGGCGAGTTGCTGTATATTT 1320
QY      441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460
DB      1321 TTTGACATATTCCTGAGATTAATCTGAGAGAGAGCCATCTCAATCTTCAGCCCTTAATC 1380
QY      461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheIys 480
DB      1381 TTCTACCTGTGTTATCTCTGACAGATGTATATCAATCAAGAGTTCCCATTTAAA 1440
QY      481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
DB      1441 ACTCTCTCAATGTTACTCATTTCTTTCACCAACATTTGTTCTTATCTAGCCAGATAT 1500
QY      501 LeuPheGluSerGlyThrLeuProProIysLeuAspValPheAspAlaValAlaArg 520
DB      1501 CTATTTGAAATGGAACTTCCCTCCCAAAATAGATGATTTGATGAGCTGTTGCCAAG 1560
QY      521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleIysLeuAsn 540
DB      1561 CACGTGAAGAGAACATGAGACAGACATCTTATGTCAGAAATGAATAATCAATTAAT 1620
QY      541 GluLeuAlaProValIysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB      1621 GAACCTGCACCTGTGAACCTCGGAGAGCCTAACCTCACTTCACTTCAACCAATAAG 1680
QY      561 GluAlaLeuLeuAspValAspSerSerProGluIysSerGlyThrGluAspAsnLeuGln 580
DB      1681 GAGGCGCTCTCTGATGTTGATTTCCAGTCCGAGGGGCTCGGAGACTGAAGATATTTACA 1740

RESULT 4
US-10-724-806-23
; Sequence 23, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:0080US
; CURRENT APPLICATION NUMBER: US/10/724,806
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; PRIOR FILING DATE: 2001-07-23
```

```
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-10-724-806-23

Alignment Scores:
Pred. No.: 2,33e-300 Length: 1743
Score: 2993.00 Matches: 580
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0

US-10-724-806-4 (1-580) x US-10-724-806-23 (1-1743)

QY 1 MetProPhHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
Db 1 ATGCTTTTCCATGTGAAGAGCTGTGATGCTATTAATCTCTTCACTTCAATATTTCTG 60
QY 21 ValGlyIleTyrAlaIleTyrPheTyrLeuSerGlyValProGluGluArgSerGlu 40
Db 61 GTTGGAATATGGGCTGCTGATGGAACCAAAACAGCGGCAACCCAGAGAGCCAGTGA 120
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db 121 GCATCATAGTCGGGGGCGGTGACATTTGTTGTGGTGGTGGTTTACCATGACAGCC 180
QY 61 ThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db 181 ACCTGGGTGGAGAGGAGCTACATCAATGGGACAGAGACATGATGGGCGAGGTTGT 240
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db 241 GGTCTAGCTTGGGCTATGCAATGCCATTTGATATCTCTGATGCTTAATTTTAAAGTGG 300
QY 101 PhePheAlaIlePhePheMetArgSerIleGlyTyrValThrMetLeuAspPropheGly 120
Db 301 TTTTGGCGAAACCTATGCGTCCAGAGGATATGTGATGCTATGTAAGCCATTCAG 360
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyMetPhe 140
Db 361 ATCTATGGAAGAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGGAGATG 420
QY 141 TyrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db 421 TGGGCTGCGAAGCAATTTTCTCTGCAATTAAGGGCCACCATTCAGCTGATCATTA 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db 481 GTGCAATATCGGTATGCTCTGCACTCATTTGCACTTTTATACCCAGAGGGG 540
QY 181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTyr 200
Db 541 CTCTACTCTGTGGCAATATCATGATGTTCCAGCTATTCGCAATTTTAAAGACTGTG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db 601 ATCAGAGTCCCTTTTCCCTGCTGCAATCCGCACTGACGAGTCAACGATTCAGAG 660
QY 221 HisAlaIleTyrLeuLeuSerProTyrPheGlyThrIleGluSerValGluValTyrThrTyr 240
Db 661 CAGCTAAATACCAAGAGTCCCTGCTGGGAAACATTCAGTGAAGCTACACCTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGluAlaTyrPheGluArg 260
Db 721 CTTGATATTTTCTGTTATTTGATGCTGGGTGGAATCCCATGCGAAGCTTCTCCAGAG 780
```

```
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGluValLeuSerPheLeuAlaIlePheGly 280
Db 781 GTCTCTCTTCAATCCACAGCACTTATGCTACAGTACTGCTTCTCCGAGCTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300
Db 841 TGCCTGTGATGGCTCTTACCCGCAATATGCAATGAGGCTATTGGACCTTCCAGACTGG 900
QY 301 AsnGluThrAlaTyrGlyTyrProAspProLysThrLysGlyGluIleAlaAspMetIleLeu 320
Db 901 AACCAAGCTGCTTACCGGATTCAGATCCCAAGACTAAGAGGAGACAGACATGATTTCTC 960
QY 321 ProIleValLeuGluIleTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db 961 CGATGCTTGTGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db 1021 TCAGCTGCTGTCAATGCTTCCAGCTGACTGCTCATCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 361 ArgAsnIleTyrGluLeuSerPheArgGluAsnAlaSerAspLysGluIleValTyrVal 380
Db 1081 CGGAATATCTACCACTTCTTCTTCAACAAATGCAATCAGACAAAGAAATGTGTGGCTC 1140
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db 1141 ATGAGATCATCTGTGCTTGTGTTGGAGCATGCAACAGCCATGGCTTGTGTCAGCAAG 1200
QY 401 ThrValTyrGlyLeuTyrPheTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db 1201 ACTGTGTATGGGCTCTGATGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 421 LeuLeuCysValLeuPheIleLysGlyTyrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db 1261 CTGCTCTGTATCTTCAATCAAGAGAACCACTTAATGGGCACTTGTGCTGTTATTT 1320
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGluIleProLeuIle 460
Db 1321 TTTGGACTATCTCTGGAATTTACTGAGGAGGACCATCTATATCTTGAAGCCCTTAATC 1380
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGluArgPheProPheLys 480
Db 1381 TTCTACCTGCTTATTAATCTTCAAGAAATGTGATATACATCAGAGGTTCCCATTTAA 1440
QY 481 ThrLeuSerMetValHisSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyr 500
Db 1441 ACTCTCTCAATGATTAACCTCAATCTTACCAACATTTGTTCTTAATCAGCAAGTAT 1500
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
Db 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGTTATGCTGTGTGCGAGAG 1560
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db 1561 CACAGTGAAGAGAACTGACAAAGACCATTTCTGTGCAAAATGAAATATCAATTAAT 1620
QY 541 GlnLeuAlaProValLysProAlaLysProArgIleThrLeuSerSerThrPheThrAsnLys 560
Db 1621 GAACTTGACCTGTGAACCTCGGAGAGCTTACCTCACTTCACTTCACTTCACTTCACT 1680
QY 561 GlnAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAsnLeuGln 580
Db 1681 GAGGCCCTCTGATGTTGATTCAGTCCGAGAGGGGCTGCGAGCTGAAGTATTTACAA 1740

RESULT 5
US-09-911-077A-5
; Sequence 5, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: PERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
```

```
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4904
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (224)..(1966)
; US-09-911-077A-5

Alignment Scores:
Pred. No.: 7.87e-296 Length: 4904
Score: 2956.00 Matches: 572
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.6% Mismatches: 4
Query Match: 98.8% Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-5 (1-4904)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20
DB 224 ATGCTTTTCATGTAGAAAGAGCTAGTACGCGATTATCTGTCTTACCTCTTAAATTTCG 283
QY 21 ValGlyIleTPrAlaAlaTPrIysThrIysAsnSerGlyAsnProGluGluArgSerGlu 40
DB 284 GTTGGAAATATGGGCTGCATGCAAGAAAACCAAAACAGGGTAAATGCAAGAACAGCCGAA 343
QY 41 AlaIleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPheThrMetThrAla 60
DB 344 GCATCATATATGGGGCCGAGACATTTGGTTGGTTGGTTTACCATGACAGCC 403
QY 61 ThrTPrValGlyGlyTyrTyrIleasnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
DB 404 ACCTGGGTGGAGGAGGTATCATCAACGGGACAGCTGAAGCAATTATGGCCAGGTTGT 463
QY 81 GlyLeuAlaTPrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
DB 464 GGTTCAGCTTGGGCTCAGGACCCCATTTGATATCTCTGATGCTGATTTTAAAGTGGCTCG 523
QY 101 PhePheAlaLysProMetArgSerIysGlyTyrValThrMetLeuAspProPheIysGln 120
DB 524 TTTTTCGAAACCTATGCGTTCAGAGGATATGTACTATGTATGACCCGTTTCAACAG 583
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
DB 584 ATCTATGGAACCGCATGGGTGGCTGCTTCATCTCCGACATGATGGAGAGATGTTTC 643
QY 141 TPrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
DB 644 TGGGCTGCAGCAATTTCTCTGCATTAAGGGCTACCATCAGCGTATCATTTGATGGAT 703
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
DB 704 GGAAACATATCGGTCAATGTCTCCGCACTCATTTGCCATCTTATATCCCTCGGGAGGG 763
QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTPr 200
DB 764 CTCTACTCTGTGGCAATATCTGATGTGTGACAGCTATCTGCATTTTAAAGATTGTGG 823
QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyPheThrAlaVal 220
DB 824 ATCAAGTCTCCATTTGGCTCTGTACATCTCTGCAGTACCCGACATTGGATTCTACGCTGTG 883
QY 221 HisAlaLysTyrGlnSerProTPrLeuGlyThrIleGlnSerValGluValTyrThrTPr 240
DB 884 CATGCTAAATATCACAGAGTCCCTGGCTGGGAAACCATTTGAATCAGTTGAAGTCTACACTGG 943
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTPrGlnAlaTyrPheGlnArg 260
```

```
DB 944 CTGATTAATTTTCGTTGTATAGTGGGTGAGATACATGCAAGCTTACTTCAGAGG 1003
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
DB 1004 GTCTCTCTTCATGCTCAGCGACCTATGCTCAGGGTGTCTCTCTCTGGACGCTTTGGG 1063
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTPr 300
DB 1064 TGCCCTGGATGGCTTACAGCCCATTTGCAATGGGGCATTTGGAGCTTCACACACTGG 1123
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIysThrIysGluGluAlaAspMetIleLeu 320
DB 1124 AACCAATGCAATATGGGTTTCCAGATCCCAAGACCAAGAGAGAGACATGATTTCTC 1183
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
DB 1184 CCGATTGTTCACAGTACCTCGCCCTGTATACATTTCTCTTGGGCTGTGGTCTGT 1243
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerPheAla 360
DB 1244 TCTGCTGCTGCATATGCTCGGCTGACATTCATTCATTCATTCAGCAAGTTCCATGTTGCT 1303
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTPrVal 380
DB 1304 CGGAATATCTACAGCTTTCCTTCAGACAAATGCAATGACACAGAAATTCGTGGGCTC 1363
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIys 400
DB 1364 ATGAGGATCATCTGTGTTGTTGTTGAGACATCTGCACAAAGCCATGGCTGTGCAGAA 1423
QY 401 ThrValIlyrGlyLeuTPrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
DB 1424 ACTGTTATGGGCTGTGGACTGAGCTGACCTTGCTCATATCATATCTTCCACAG 1483
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1484 CTGCTCTGTGATCTTTTCATCAAGAAACCAACATTTAGGGGACGTTGCTGTTATAT 1543
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyProTyrLeuTyrLeuGlnProLeuIle 460
DB 1544 TTTGACCTTTTCCTGAGATATCCGAGAGAGCCATATCTTATCTTGACGCCCTTAATC 1603
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
DB 1604 TTCTACCTCGTTATTTACCTGACAAAGATGATATATCAATCAGAGGTTCCCATTTAAA 1663
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
DB 1664 ACTCTCTCAGATGTTACCTCATTTCTTTACCAACATTTTGTTTCTTATCTACCAAGTAT 1723
QY 501 LeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaArg 520
DB 1724 CTATTTGAAAGTGAACCTTGCTCCAAAATTAAGATATTTGATGCTGTGTCTCAAGG 1783
QY 521 HisSerGluGluAsnMetAspLysThrIleLeuValAlaGlnGluAsnIleLysLeuAsn 540
DB 1784 CACAGTGAAGAGAAATGACAAACATTTCAAGTCAAGAAAGAAACATCAAAATTAAT 1843
QY 541 GlyLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1844 GAACCTTGCACTGTAAAGCTTCGACAGAGCTTACCTCAGTTCAACTTTCACCAATAAA 1903
QY 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580
DB 1904 GAGGCTCTCTGATGTTGATTTCAAGTCCAGAGGGATCTGGGACATGAGATTAATCAACA 1963

RESULT 6
US-10-724-806-5
; Sequence 5, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
```



APPLICANT: APPASUNDARAM, SUBRAMANIAM  
 APPLICANT: FERGUSON, SHAWN  
 TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
 FILE REFERENCE: VBLT:008US  
 CURRENT APPLICATION NUMBER: US/10/724,806  
 PRIOR FILING DATE: 2003-12-01  
 PRIOR APPLICATION NUMBER: US/09/911,077A  
 NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 4904  
 TYPE: DNA  
 ORGANISM: Rattus norvegicus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (224)..(1966)  
 US-10-724-806-5

## Alignment Scores:

Pred. No.:	7,87e-296	Length:	4904
Score:	2956.00	Matches:	572
Percent Similarity:	99.3%	Conservative:	4
Best Local Similarity:	98.6%	Mismatches:	4
Query Match:	98.8%	Indels:	0
		Gaps:	0

US-10-724-806-4 (1-580) x US-10-724-806-5 (1-4904)

```

QY      1 MetProphHisValGluGlyLeuValAlaIleIleuPheTyrIleuLeuIlePheLeu 20
Db      224 ATGCTTTCATGTACAGAGAGCTAGTAGCATTAATCTCTTCTTAATATTCTG 283

QY      21 ValGlyIleTyrAlaIleTyrIleuValAsnSerGlyAsnProGluGluAsnSerGlu 40
Db      284 GTTGAATATGGGCTCATGGAACCAAAACAGGATATCAGAGAACGACGAGAA 343

QY      41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60
Db      344 GCCATCATGTTGGGGCCGAGACATTGGTTGTTGGTTGGTTTACATGACAGACC 403

QY      61 ThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db      404 ACCTGGTGTGAGAGAGTTACATCAACGGGACAGCTGAGACAGTTATGGCCAGGTTGT 463

QY      81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerIleuSerIleuLeuGlyIleu 100
Db      464 GGTCTAGCTTGGGCTCAGGACCCATGATGATTTCTGTGATCTGATTTTAGGGCTTG 523

QY      101 PhePheAlaIlePheMetArgSerIleGlyTyrValThrMetLeuAspProPheLeuGln 120
Db      524 TTTTTCGAAAACCTATGCGTTCCAGGATATGTGACTATGTAGACCCGTTTCAACAG 583

QY      121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe 140
Db      584 ACTATGAAAGCGCATGGGTGGGCTGCTGCTTCATCCCTGACATGATGGAGAGATGTT 643

QY      141 TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db      644 TGGGCTGACGCAATTTTCTCTGATTAGGGGCTTACATCAGCGTAATCATGTATGGAGT 703

QY      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleuTyrThrIleuValGlyGly 180
Db      704 GTGAACATATCGGTCTATGCTCCGACATCATTTTATACCCCTCGGGAGAGG 763

QY      181 LeuTyrSerValAlaTyrThrAspValValGluLeuPheCysIlePheIleGlyLeuTrr 200
Db      764 CTTCTACTCTGTGCAATATCTGATGTGTACAGCTATTCTGCAATTTTATAGGATGTGG 823

QY      201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db      824 ATCAGTGTCCCATTTGGCTGTGACATCTGACAGTACCGACATTTGATTCACCTGCTGTG 883
  
```

```

QY      221 HisAlaIleTyrGlnSerProTrrIleuGlyThrIleGluSerValGluValTyrThrTrp 240
Db      884 CATGTAAATATACAGAGTCCCTGGCTGGGACCATGATGATGATGATGATGATGATGATGAT 943

QY      241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrrGlnAlaTyrPheGlnArg 260
Db      944 CTTGATTAATTTTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1003

QY      261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      1004 GTCTCTCTTCATTCGTCAGACGACCTATGCTCAGTGTCTCTCTCTGAGGCTTTGGG 1063

QY      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrr 300
Db      1064 TGCCTGATGATGCTTACAGACCATTTGATGGGCTTGGAGCCCTCCACAGACTGG 1123

QY      301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMetIleu 320
Db      1124 AACCAACTGCATATGGGTTTCCAGATCCCAAGACCAAGAGAGAACAGACATGATTTCTC 1183

QY      321 ProIleValLeuGlnTyrIleuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db      1184 CCGATGTTCTTACAGTACCTCTGCTGTACATTTCTTTGGGCTTGGTGTGCTGTT 1243

QY      341 SerAlaAlaValMetSerSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db      1244 TCTGCTGTGTATGCTCTCGGCTGATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1303

QY      361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleAlaTrrVal 380
Db      1304 CGGAATATCTACACGCTTCTTCCCTTCCAGCAAAATGATCAGACAGCAAAATGATGATGATGATGAT 1363

QY      381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
Db      1364 ATGAGATATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1423

QY      401 ThrValTyrGlyLeuTrrTyrIleuSerSerAspLeuValTyrIleIlePheProGln 420
Db      1424 ACTGTATATGGGCTGTGATCTGACGACCTGCTGATCAATCAATCAATCAATCAATCAATCAATCA 1483

QY      421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db      1484 CTGCTCTGTGATCTTCTTCAAAAGAACCAACATTAATGGGCAATGTGTGTATAT 1543

QY      441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrIleuGlnProLeuIle 460
Db      1544 TTTGACCTTTCTTGGAATTTACCGAGAGAGACCATATCTATCTTGACAGCCCTTATATC 1603

QY      461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db      1604 TTCTACCTCGTTATTAACCTGACAGAAATGATATATACATCAGAGGTTCCCATTTTAA 1663

QY      481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrIleuAlaLysTyr 500
Db      1664 ACTCTCTCATAGGTATACCTCATTTCTTACCAATTTGGTTCCATCTAGGCAAGTAT 1723

QY      501 LeuPheGluSerGlyThrIleuProLysIleAspValPheAspAlaValAlaIleArg 520
Db      1724 CTATTTGAAAGTGAACCTTGCTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1783

QY      521 HisSerGluGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
Db      1784 CACAGTGAAGAGAACATGACAAACATCTAGTCAGAAATGAAATCAATCAATTAATTAAT 1843

QY      541 GluLeuAlaProValLysProArgGlnSerIleuThrLeuSerSerThrPheThrAsnLys 560
Db      1844 GAACTTGACCTGTAAAGCTCGACAGACCTTAACCTCAGTTCAACTTTCACCAATTAAT 1903

QY      561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspLeuLeuGln 580
Db      1904 GAGGCTCTCTTGATGTATTTTCAATCCAGAGGAGATCTGGAGCTGAAGATTAATTAATTAAT 1963
  
```

RESULT 7

```
US-09-911-077A-1
; Sequence 1, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPASUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLF:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-1

Alignment Scores:
Pred. No.: 9,496-280 Length: 1743
Score: 2795.00 Matches: 537
Percent Similarity: 96.6% Conservative: 23
Best Local Similarity: 92.6% Mismatches: 20
Query Match: 93.4% Indels: 0
DB: Gaps: 0

US-10-724-806-4 (1-580) x US-09-911-077A-1 (1-1743)

QY 1 MetProPheHisValGIuGIyLeuValAlaIleIleuPheTyrlEuLeuIlePheLeu 20
DB 1 ATGGCTTCCAGTGTGAAGAGCTGATAGCTATCATCTGCTTACCTTCTAATTTTGGCTG 60
QY 21 ValGIyIleTPrAlAlaIleTPrLysThrLysAsnSerGIyAsnProGIuGIuArgSerGIu 40
DB 61 GTTGGAAATATGGGCGCTCGGCAACCAAAACAGTGGCGCGCAAGAGCGACAGCGAA 120
QY 41 AlaIleIleValAlGIyGIyArgAspIleGIyLeuLeuValAlGIyPheThrMetThrala 60
DB 121 GCCATCATATGTTGGTGGCCGAGATATGGTTATTGTTGGTGGATTTTACATGACAGCT 180
QY 61 ThrTPrValAlGIyGIyTyrlIleAsnGIyThrAlaGIuAlaValTyrgIyProGIyCys 80
DB 181 ACCTGGGTCGGAGGAGGTATATCAATGGCACAGCTGAAGCAGTTTATGTACAGGTTAT 240
QY 81 GIyLeuAlaTPrAlaHisAlaProIleGIyTyrlSerLeuSerLeuIleLeuGIyGIyLeu 100
DB 241 GGCTTAGCTTGGGCTCAGGACCAATTTGSAATTTCTTAGTCTGATTTTAAAGTGGCTTG 300
QY 101 PhePheAlaLysProMetArgSerLysGIyTyrlValThrMetLeuAspProPheLysGIu 120
DB 301 TTCTTTGCAAAACCTATATGCTTCAAAAGGGGTATGTGACATGTTTGAACCGCTTTACGAA 360
QY 121 IleTyrgIyLysArgMetGIyGIyLeuLeuPheIleProAlaLeuMetGIyGIuMetPhe 140
DB 361 ATCTATGGAAAAACGATGGCGGACCTCTGTTTATTTCCGCACTGATGGAGAAATGTTC 420
QY 141 TrpAlaIleAlaIlePheSerAlaLeuGIyAlaThrIleSerValIleIleAspValAsp 160
DB 421 TGGGCTGGACAAATTTCTCTGCTTTGGAGCGCACCATGAGCGTGAATCAATGAGGAT 480
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrlThrLeuValAlGIyGIy 180
DB 481 ATGCACATTTCTGTCATCATCTCTGCACTCATGCACTCTGTACACACACTGGGAGGAGG 540
QY 181 LeuTyrlSerValAlaTyrlThrAspValValAlGIuLeuPheCysIlePheIleGIyLeuTPr 200
DB 541 CTCTATTTCTGTGGCCTTACACTGATGTCTGTGAGCTTTTGGCAATTTTGTAGGGCTGTGG 600
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGIyPheThrAlaVal 220
```

```
DB 601 ATCAGCGTCCCTTTGCAATGTGCATCTCGAGTCGACAGACATGGGTTCACTCTGTG 660
QY 221 HisAlaLysTyrgIuInSerProTPrPheGIyThrIleGIuSerValGIuValTyrlThrTPr 240
DB 661 CATGCCAATATCCAAAGCCGCTGGGAACTGTGACTCATCTGAAGTCACTACTCTTGG 720
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGIyGIyIleProTPrGIuAlaTyrlPheGIuArg 260
DB 721 CTTGATATGTTTCTGTGTGTGATGTGGGTGAATCCCATGGCAAGCATCTTTCAGAGG 780
QY 261 ValIleuSerSerSerSerAlaThrTyrlaGIuValIleSerPheLeuAlaIlePheGIy 280
DB 781 GTTCTCTCTTCTCTCTGACGACCTATGCTCAAGTCTGCTCTCTCTGCACTTTGGG 840
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGIyAlaIleGIyAlaSerThrAspTPr 300
DB 841 TGCCCTGGATGGCCATCCAGCCCATCTCATCTGGGGCATTGGAGCATCAACAGACTGG 900
QY 301 AsnGIuThrAlaTyrgIyTyrlProAspProLysThrLysGIuGIuAlaAspMetIleLeu 320
DB 901 AACCACTGCATATAGGGCTTCCAGATCCCAAGACTCAAGAGAGAGACATGATTTTGA 960
QY 321 ProIleValIleuGIuInTyrlLeuCysProValTyrlIleSerPhePheGIyLeuGIyAlaVal 340
DB 961 CCAATTTGTTCTGCACTATCTCTGCTGTATATTTCTTTCTTTGGTCTGTGTCAGATT 1020
QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
DB 1021 TCTGCTGCTGTTATGTATCATACAGCAAGATTTCTTCATCTGTGACGAATGTCATGTTTGA 1080
QY 361 ArgAsnIleTyrgIuInLeuSerPheArgGIuAsnAlaSerAspLysGIuIleValTPrVal 380
DB 1081 CGGAACATCTTACACACTTCTCTTCCAGCAAAATGCTTGGCAAAAGAAATCTTTGGGTT 1140
QY 381 MetArgIleThrValLeuValPheGIyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1141 ATGCATATCAACAGTGTGTTGTTGGACATCTGCACACCATGCTGCTGTGACGAA 1200
QY 401 ThrValTyrgIyLeuTPrTyrlLeuSerSerAspLeuValTyrlIleIlePheProGIu 420
DB 1201 ACTGTGATAGGCTGTGGTATCCTCAGTCTGACCTTGTTTACATGTTATCTTCCCCGAG 1260
QY 421 LeuLeuCysValLeuPheIleLysGIyThrAsnThrTyrgIyAlaValAlaGIyTyrlIle 440
DB 1261 CTGCTTGTGTACTCTTTGTTAAGGAAACCAACACTTATGGGGCGGTGGCAGATTATGT 1320
QY 441 PheGIyLeuPheLeuArgIleThrGIyGIuProTyrlLeuTyrlLeuGIuInProLeuIle 460
DB 1321 TCTGGCTCTTCTCTGAGAAATATCTGAGGGAGGACATATCTGTATCTTCAGCCCTTGATC 1380
QY 461 PheTyrlProGIyTyrlTyrlSerAspLysAsnGIyIleTyrlAsnGIuArgPheProPheLys 480
DB 1381 TTCTACCTCGGCTATTAACCTGATATATGATATATATATCAAGAAATTTCCATTTAA 1440
QY 481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrlLeuAlaLysTyrl 500
DB 1441 ACACTTGCCATGTATCATCTTTTAAACCAACATTTGCACTCTCTATCTTACCAAGATAT 1500
QY 501 LeuPheGIuSerGIyThrLeuProProLysLeuAspValPheAspAlaValAlaIleArg 520
DB 1501 CTATTTGAAAGTGAACCTTGGCACTTAATTAAGTATTTGAAGCTCTTCTTCCAGAA 1560
QY 521 HisSerGIuGIuAsnMetAspLysThrIleLeuValArgAsnGIuAsnIleLysLeuAsn 540
DB 1561 CACAGTGAAGAAAAATAGATATAGCAATTTCTGTCAAAAAATATTAATTTAGATAT 1620
QY 541 GIuLeuAlaProValLysProArgGIuInSerLeuThrLeuSerSerThrPheThrAsnLys 560
DB 1621 GAACCTTGCACTTGTAAACCGACAGACAGCATGACCTTCAAGCTTCAACCAATTA 1680
QY 561 GIuAlaLeuLeuAspValAspSerSerProGIuGIySerGIyThrGIuAspAsnLeuGIu 580
```

Db 1661 GAGGCTTCCTGATGTTGATTCAGTCCAGAGGGTCTGGAGCTGAGATTAATTACAG 1740  
RESULT 8  
US-10-724-806-1  
; Sequence 1, Application US/10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
; FILE REFERENCE: VBLT:00805  
; CURRENT APPLICATION NUMBER: US/10/724,806  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US/09/911,077A  
; PRIOR FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1743  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1743)  
US-10-724-806-1

## Alignment Scores:

Pred. No.:	9.49e-280	Length:	1743
Score:	2795.00	Matches:	537
Percent Similarity:	96.6%	Conservative:	23
Best Local Similarity:	92.6%	Mismatches:	20
Query Match:	93.4%	Indels:	0
DB:	9	Gaps:	0

US-10-724-806-4 (1-580) x US-10-724-806-1 (1-1743)

QY 1 MetProPheHisValGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
Db 1 ATGGCTTTCATGTTGAGAGACTGATGATCATCGTTCCTTACTTAAATTTTGGCG 60  
QY 21 ValGlyIleTyrAlaIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIleTyrIle 40  
Db 61 GTTGAATATGAGGCTGCTGAGAACCAAAACAGTGCAGCGCAGAGAGCGACGAA 120  
QY 41 AlaIleIleValGlyLeuArgAspIleGlyLeuLeuValGlyIlePheThrIleAla 60  
Db 121 GCCATCATAGTGTGGCCGAGATATGGTTTATGGTGTGGATTACCATGACAGCT 180  
QY 61 ThrTyrValGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
Db 181 ACCTGGCTCGAGAGGATATCATAGGACACGCTGAGAGCGTTTATGATCAGGTTAT 240  
QY 81 GlyLeuAlaTyrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeu 100  
Db 241 GGCTAGCTGGGCTCAGGACCAATGGATATTCCTTACTGATTTTAACTGGGCTCG 300  
QY 101 PhePheAlaIlePheSerAlaSerIleGlyTyrValThrMetLeuAspProHelysGln 120  
Db 301 TTTCTTGGAAAACCTATGGTTCAAAGGGGTATGTGACCATGTTAGACCCGTTTCAGCA 360  
QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPhe 140  
Db 361 ATCTATGGAAAACGATGGGCGGACTCCGTTTATTCCTGCACTGATGGAGAAAATGTTTC 420  
QY 141 TTPAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
Db 421 TGGGCTGACGAAATTTCTCTGCTTGGAGGACCATCAGCGTGTATCATCATGATGGAT 480  
QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
Db 481 ATGCACATTTCTGTATCATCTCTGTGACATCATTTGCCACTCTGTACACACTGGTGGAGGG 540

QY 181 LeuTyrSerValAlaTyrThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTyr 200  
Db 541 CTCATATTCGTGGCTTACACATGATGCGTTCAGCTCTTTGATTTTGTAGGGCTGTGG 600  
QY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220  
Db 601 ATCAGCGTCCCTTGGATTCATGTCATCTCGAGTCGACATCGGGTTCACCTGCTGG 660  
QY 221 HisAlaIleTyrGlnSerProTyrLeuGlyThrIleGlnSerValGluValTyrThrTyr 240  
Db 661 CATGCCAANTACCAAAAGCCGTGGGACTGTTGACTCTCTGAACTCTGAACTCTACTTGG 720  
QY 241 LeuAspAsnPheLeuLeuMetLeuGlyIleProTyrGlnAlaTyrPheGlnArg 260  
Db 721 CTGATATGTTTCTGTGTTGATGCTGGGTGAATCCATGGCAAGCATCTTTCAGAG 780  
QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
Db 781 GTTCTCTCTTCTTCTTCCAGCCACTATGCTCAAGTGTCTCTCTCTGCGACCTTTCGGG 840  
QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyValAspThrAspTyr 300  
Db 841 TGCCTGGATGCGCATCCAGCCATCTCATTTGGGCGCATTTGAGACATCAACAGACTGG 900  
QY 301 AsnGlnThrAlaTyrGlyTyrProAspProIleTyrIleGlyGluAlaAspMetIleLeu 320  
Db 901 AACCAAGCTGATATGGGCTTCAGATCCCAAGACTACAGAGAGGACAGCATGATTTTA 960  
QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
Db 961 CCATATGTTCTCGACATATCTTCCCTCGTATATTTCTTTTGTGCTTGGTGGTACAGTT 1020  
QY 341 SerAlaAlaValMetSerSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
Db 1021 TCTGCTGCTGTTATGATCATCAGAGATTTCTTCACTTGTCTGCAAGTTCCATGTTTGA 1080  
QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTyrVal 380  
Db 1081 CGGAACATCTACACACTTTCCTTCAAGCAAAATGCTTGGAGCAAAAGAAATCGTTGGGTT 1140  
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400  
Db 1141 ATGCCAATCACAGTGTGTTGTTGTGGAGCATCTGCAGACGCCATGGCTTGTCTGACGAA 1200  
QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
Db 1201 ACTGTGTAGGGCTGTGACTCAGTCTGACCTTGTATTAATGATGTTATCTTCCCCAG 1260  
QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaIleGlyTyrIle 440  
Db 1261 CTGCTTTGTGATCTTTGTTTAAGGAACCAACACTGATGGGCGGTGACAGTTATGTT 1320  
QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIle 460  
Db 1321 TCTGGCTCTTCTCTGAGATTAAGTGAAGGGGAGCAATCTGTATCTTACGCTTGAATC 1380  
QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProHelys 480  
Db 1381 TTTTACCTCGGCTATTTACCTGATGATATGATATATATATATATATATATATATATAT 1440  
QY 481 ThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
Db 1441 ACACTTGGCATGATACATCATCTTCAACCAACATTTGATGATCTCCATCTTACCAAGTAT 1500  
QY 501 LeuPheGlnSerGlyThrLeuProProLysLeuAspValPheAspAlaValAlaLys 520  
Db 1501 CTATTTGAAAGGAAACCTTGCACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1560  
QY 521 HisSerGlnGluAspMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
Db 1561 CACAGTGAAGAAACATGATGATGACATTTCTGTCAAAATGAAATATTAATTAATTAAT 1620  
QY 541 GluLeuAlaProValLysProArgLysSerLeuThrLeuSerSerThrPheThrAsnLys 560

```
Db      1621 GAACCTTGACCTTGACACCGACAGACGACCGCTCAGCTCAACTTCCACCAATAAA 1680
Qy      561 GUAUAlaLeuAspValAspSerProGluGlySerGlyThrGluAspAsnLeuGln 580
Db      1661 GAGGCTTCCTTGAGTGTGATTCAGTCCAGAGAGGCTCGGACCTGAGATTAATTACAG 1740

RESULT 9
US-09-911-077A-9
; Sequence 9, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1761)
US-09-911-077A-9

Alignment Scores:
Pred. No.:      1,01e-279      Length:      1813
Score:          2795.00        Matches:     537
Percent Similarity: 96.6%      Conservative: 23
Best Local Similarity: 92.6%    Mismatches:  20
Query Match:     93.4%         Indels:       0
DB:              3            Gaps:          0

US-10-724-806-4 (1-580) x US-09-911-077A-9 (1-1813)
Qy      1 MetProPheHisValGlyLeuValAlaIleIleLeuPheThrLeuLeuIlePheLeu 20
Db      19 ATGGCTTTCATGCTGGAAGGAGCTGATGCTATCATGCTTACCTTTAATTTTGGCG 78
Qy      21 ValGlyIleThrAlaAlaIleThrIleAsnSerGlyAsnProGluGlnArgSerGlu 40
Db      79 GTTGAAATATGGGCTGCTCGGAGAACCAAAAACAGTGGCGACGAGAACGCGACGGA 138
Qy      41 AlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyPheThrMetThrAla 60
Db      139 GCCATCATATGTTGGTGGCCGAGATATGCTTATTTGGTTGGATTTTACATGACAGCT 198
Qy      61 ThrTrpValGlyGlyGlyThrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80
Db      199 ACCTGGGTCGGAGGAGGGGATATCATGACAGCGACGAGGATTATATACAGGTTAT 258
Qy      81 GlyLeuAlaIleThrAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100
Db      259 GGCCTAGCTTGGGCTCGAGCACCAATGGATATTCCTTAGTCTGATTTTAAAGTGGCCCG 318
Qy      101 PhePheAlaLysProMetArgSerIleGlyTyrValIleThrMetLeuAspProPheLysGln 120
Db      319 TTCTTTGCAAAACCTATGCTTCAAAAGGGGATATGAGCCATGTATGACCCGTTTCAGCAA 378
Qy      121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlnMetPhe 140
Db      379 ATCTATGGAAGAACGATGGGCGGACCTCGTTTATTTCCGCACTGATGGGAGAAATGTTTC 438
Qy      141 TrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
Db      439 TGGGCTGACGAAATTTCTCTGCTTTGGAGGACCATCAGCGTGTATCATCATGTGGAT 498
Qy      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180

Db      499 ATGCACTTTCTGTCATATCTCTGCACTCATGCTCACTGTATACACTGTTGGAGGG 558
Qy      181 LeuTyrSerValAlaIleThrAspValAlaGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      559 CTCTATTTCTGAGGCTTACAGTATGCTTCAAGCTCTTTTGATTTTGTAGGGCTGTAG 618
Qy      201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db      619 ATCAGCGTCCCTTTGCAATGTCATCTCCAGTCGACGAGCATGCGGGTTCATCTCTGTG 678
Qy      221 HisAlaLysTyrGlnSerProTrpLeuGlyThrIleGlySerValGlyValTyrThrTrp 240
Db      679 CATGCCAAATACCAAAACCGGTGGGAACTGTGACTCATCTGAAAGTCTACTCTTGG 738
Qy      241 LeuAspAsnPheLeuLeuMetLeuGlyGlyIleProTrpGlnAlaIlePheGlnArg 260
Db      739 CTGTATGATTTTCTGTGTGTGATGCTGGGTGAATCCCATGCAAGCATCTTCCAGAG 798
Qy      261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280
Db      799 GTTCTCTCTTCTCTCCAGCCACCTATGCTCAAGTGTGCTTCTCTGCAAGCTTTCCGG 858
Qy      281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrp 300
Db      859 TGCTGTGATGAGCCATCCAGCCATCTCATTTGGGGCCATTGGAGCATCAACAGACTGG 918
Qy      301 AsnGlnThrAlaTyrGlyTyrProAspProIleThrIleSerPhePheGlyLeuGlyAlaVal 320
Db      919 AACCAAGCTGCAATATGGGCTTCCAGATCCCAAGACTACAGAAAGAGCAGACATGATTTA 978
Qy      321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340
Db      979 CCAAATGTTCTGCACTATCTGCTGCTGTGATATTTCTTCTTGGTCTGTGGTCCAGTT 1038
Qy      341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360
Db      1039 TCTGCTGCTGTATATGTCATACAGCAATCTTCCATCTTGTACAGCAAGTTCATGTTTGA 1098
Qy      361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTrpVal 380
Db      1099 CGGAACATCTACACACTTCTCTTCAGCAAAATGCTTGGCAAAAGAAATCGTTGGGTT 1158
Qy      381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle 400
Db      1159 ATGCCAAATCACAGTGTGTTGTGTGGACATCTGCAACAGCCATGCGCTGTGACGAAA 1218
Qy      401 ThrValTyrGlyLeuTrpTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420
Db      1219 ACTGTGTATGGGCTGTGATCTGCTGACCTGTGACCTGTATTACATGCTTATCTTCCCCAG 1278
Qy      421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
Db      1279 CTGCTTTGTATGACTCTTGTATAGGGAACCAACACTATGGGGCGGTGGAGGTTATGTT 1338
Qy      441 PheGlyLeuPheLeuArgIleThrGlyGlyGlnProTyrLeuTyrLeuGlnProLeuIle 460
Db      1339 TCTGGCTCTTCTCGAAGATTAACGGAAGGGAGCCATATCTGTATCTTCAGGCTTGTATC 1398
Qy      461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480
Db      1399 TTCTACCTCGGCTATTAACCTGATGATATATGATATACAAATTTCCATTATAA 1458
Qy      481 ThrLeuSerMetValThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500
Db      1459 ACACTTGCAATGATACATATTTCTTAACCAACATTTGCATCTCTATCTAGCCAAAGTAT 1518
Qy      501 LeuPheGlnSerGlyThrLeuProProIleLysLeuAspValPheAspAlaValAlaIleArg 520
Db      1519 CTATTGGAAGAGAACCTTGGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1578
Qy      521 HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540
```



Db 1519 CTATTGAAAGGAACTTGGACCTTAATTGATTTGATGCTGTGTTCCAGAA 1578  
Qy HisSerGluGluAspMetAspLysThrIleLeuValArgAsnGluAsnIleLeuLeuAsn 540  
Db 1579 CACAGTGAAGAAAATATGATAGACATATCTCTTGCACAAAATGAAAATTTAAATTAGAT 1638  
Qy 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
Db 1639 GAACCTTGACCTTGTAAGCGACAGACAGCATGACCTCAGCTCAACTTTCACCAATAAA 1698  
Qy 561 GluAlaLeuLeuAspValAspSerSerProGluGlySerGlyThrGluAspAsnLeuGln 580  
Db 1699 GAGGCTTCTTGATGTTGATTCCAGTCCAGAGAGGCTTGAGACTGAGATTAATTACAG 1758  
RESULT 11  
US-10-241-784-1  
; Sequence 1, Application US/10241784  
; Publication No. US20040048261A1  
; GENERAL INFORMATION:  
; APPLICANT: Bayer Corporation  
; TITLE OF INVENTION: Invertebrate Choline Transporter Nucleic Acid, Polypeptides and U  
; FILE REFERENCE: Theroef  
; CURRENT FILING DATE: 2002-09-11  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1833  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1833)  
; OTHER INFORMATION:  
US-10-241-784-1  
Alignment Scores:  
Pred. No.: 4,93e-144 Length: 1833  
Score: 1491.50 Matches: 294  
Percent Similarity: 72.7% Conservative: 81  
Best Local Similarity: 57.0% Mismatches: 122  
Query Match: 49.8% Indels: 19  
Gaps: 7  
US-10-724-806-4 (1-580) x US-10-241-784-1 (1-1833)  
Qy 4 HisValGluGlyLeuValAlaIleIleLeuPheTyrrLeuLeuIlePheLeuValGlyIle 23  
Db 7 AATATCGCTGGCGGTGGAGCATCGTCTTCTTCACTCTTAATCCTGGTCTGGCATTT 66  
Qy 24 TrpAlaAlaTrpLysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle 43  
Db 67 TGGGCGGTCGCAAG---AAGCAGTCCGGCATGATTCGAG-----GAGAGGTCAG 117  
Qy 44 ValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpVal 63  
Db 118 CTGGCGGACGCTCCATCGGCTCTTCCGTGGGCATCTTACCATGACGGCCACCTGGGTG 177  
Qy 64 GlyGlyGlyTrpIleAsnGlyThrAsnGluAlaValTyrrGlyProGlyCysGlyLeuAla 83  
Db 178 GGTGGCGGCTACATCAACGCGACGCGGCGAGGCTATATAC-----ACATCGGCTGTG 231  
Qy 84 TrpAlaHisAlaProIleGlyTyrrSerLeuSerLeuIleLeuGlyGlyLeuPhePheAla 103  
Db 232 TGGTGGCAGGCTCCATTGGATACGCTCTAAGCTTGATTTGGTGGCATCTTCTTGGC 291  
Qy 104 LysProMetArgSerLysGlyTyrrValThrMetLeuAspProPheLysGlnIleTyrrGly 123  
Db 292 AATCCCATGCGCAGAGGAGGTTACATCACCATGTTGATCCGTTGACGAGATTCCTTTGGT 351  
Qy 124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPheThrAlaAla 143

Db 352 GAGCGGATGGAGAGATTCCTCTCTGCTCATCGCGGAGGATCTTTGGGCACGC 411  
Qy 144 AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle 163  
Db 412 GGCATCTCGGCTGACCTTGGCGCCACTATCGGATCATGACATGATGATCACCAGCC 471  
Qy 164 SerValIleValSerAlaLeuIleAlaIleLeuTyrrThrLeuValGlyGlyLeuTyrrSer 183  
Db 472 TCGGATGATCCGTCTCTCGATCGCATCTTCAACACACTGTTGGGAGCTGATCC 531  
Qy 184 ValAlaTyrrThrAspValValGlyLeuPheCysIlePheIleGlyLeuTrpIleSerVal 203  
Db 532 GTGGCGTATACGAGAGTATCATCGATTGTTCTGCATCTTCAATCGATCGTGTGATGTCATT 591  
Qy 204 ProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaLys 223  
Db 592 CCCTTCGCTGGAGAGACAGAGAGCTGGGCGAGCTG-----AGTGAC 633  
Qy 224 TyrrGlnSerProTrpLeuGlyThrIleGlnSerValGluValTyrrThrTrpLeuAspAsn 243  
Db 634 CTGAGCGTGAATTGGATTGGGACCTGAGCGCTAAMAAAGCATTTGCTGTACATACCTAC 693  
Qy 244 PheLeuLeuMetLeuGlyGlyIleProTrpGluAlaTyrrPheGluArgValLeuSer 263  
Db 694 GCGTGTGCTGCTGTTTGGTGGCATTTCCCTGGACAGGTCTACTTCCAGCG----- 744  
Qy 264 SerSerSerAlaThrTyrrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuVal 283  
Db 745 ---CAAAACGGCAGAGAGGCGCCAGCTTCTGCTATGTTGACGCGCGAGATCATTTTG 801  
Qy 284 MetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThr 303  
Db 802 ATGGCCATTTCCCGCTGCTCATCGAGCATTTGCCAAGCATCACCTTGAAACAGACA 861  
Qy 304 AlaTyrrGlyTyrrProAspProLysThrLysGluGluAlaAspMetIleLeuProIleVal 323  
Db 862 GATTCAAGGACCTTATCCCTGACCGTGACGAGACAGACAGATATCTGCGCCATGGTG 921  
Qy 324 LeuGlnTyrrLeuCyProValTyrrIleSerPhePheGlyLeuGlyAlaValSerAlaAla 343  
Db 922 CTGCAGTACCTAGCCCTGACCTGTCGTCCTTCTTGATTTGGCGCTTTCCCGCGCC 981  
Qy 344 ValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363  
Db 982 GTGATGTCTCTCGCCGACCTCTCGGTGCTCCCGCGCTCATGTGCTGGAAACGTG 1041  
Qy 364 TyrrGlnLeuSerPheArgGluAsnAlaSerAspLysGluIleValTrpValMetArgIle 383  
Db 1042 TACAAATGATTTTCCGTCAGAGCGTCCAGATGAAATATATTGGGTATGCGAGTTC 1101  
Qy 384 ThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThrValTyrr 403  
Db 1102 GCCATCATTTGTGGTGGCATCTGCTACATCAGTCCCTCACCATTCCTCCATCTAC 1161  
Qy 404 GlyLeuTrpTyrrLeuSerSerAspLeuValTyrrIleIleIlePheProGlnLeuLeuCys 423  
Db 1162 GGTGTGTGTGTCATGTCGTCGATCTGATCTAAGCATTTGTTTCCCGAGCTACGATGATG 1221  
Qy 424 ValLeu---PheIleLysGlyThrAsnThrTyrrGlyAlaValAlaGlyTyrrIlePheGly 442  
Db 1222 GTGGTCACTTCAAAAGACATCGAACACGTCAGGACGCTTCGCGATCACTTGTGGCC 1281  
Qy 443 LeuPheLeuArgIleThrGlyGlyGluProTyrrLeuGlnProLeuIlePheTyrr 462  
Db 1282 CTGGGCATTCGACGTGCGGCGGTGAGGCACTTGGAGCTGCGCATTTGATCAAGTAT 1341  
Qy 463 ProGlyTyrrTyrrSerAspLysAsnGlyIleTyrrAsnGlnArgPheProPheLysThrLeu 482  
Db 1342 CCGGCTACGACGAGAGAACCAAG-----GAGCAGATGTTCCCTTCCGACCAAG 1392  
Qy 483 SerMetValThrSerPhePheThrAsnIleCysValSerTyrrLeuAlaTyrrLeuPhe 502

Db 1393 GCCATGCTGCTCAGCCTGTGTCAGCTCATCTCGGTCTCCTGTGAGCACTAAATGATGTTT 1452  
Qy 503 G|USeRg|YThr|euPro|LyS|LeuAsp|ValPhe|Asp|Ala|Val 518  
Db 1453 GAGTCGGGCAAGTTGCCGCCGACTGACACTACTTCCGCTGTGTGTC 1500

RESULT 12  
US-09-911-077A-7  
; Sequence 7, Application US/09911077A  
; Publication No. US20030114399A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: AAPSUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/09/911,077A  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1985  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (9)..(1739)  
US-09-911-077A-7

Alignment Scores:  
Pred. No.: 4,34e-139 Length: 1985  
Score: 1444.50 Matches: 292  
Percent Similarity: 68.6% Conservative: 93  
Best Local Similarity: 52.0% Mismatches: 145  
Query Match: 48.3% Indels: 31  
Gaps: 9

US-10-724-806-4 (1-580) x US-09-911-077A-7 (1-1985)

Qy 7 G|YLeu|ValAla|Ile|LeuPhe|Yr|LeuLeu|IlePhe|LeuVal|G|Y|Ile|Trr|Ala|Ia 26  
Db 24 GGAATCGTGCCCATGTTGTTCTTCTACGCTCATCTTGTCTGGAAATGCGGGGT 83  
Qy 27 T|P|Y|S|Thr|Lys|Asn|Ser|G|Y|Leu|Pro|G|U|G|U|A|U|G|-----Ser|G|U|A|I|I|e 42  
Db 84 AGAAATTCGAAAGTTTCAAAAGCTTGAATTCAGAAAGCCGCGCGGACGGAAGAGGTG 143  
Qy 43 I|Leu|Al|G|Y|L|A|Y|A|S|P|I|L|E|G|Y|L|E|U|V|A|G|Y|L|P|H|E|T|H|A|L|A|T|T|P 62  
Db 144 ATGTTAGCTGGGAGAAATCGGAACCTGTTGCGAATTTTCAATGACTGCCACGTGG 203  
Qy 63 V|A|G|Y|L|G|Y|L|Y|T|I|L|E|S|N|G|Y|T|H|A|G|U|A|V|A|I|Y|G|Y|P|R|O|G|Y|C|Y|S|G|Y|L|E|U 82  
Db 204 GTTGGGGGCTTATATCATGAGAACCGCGAGCTCTGTATATGGA-----GGTCTC 257  
Qy 83 A|L|A|T|P|A|L|H|I|A|L|A|P|R|O|I|L|E|G|Y|T|Y|S|E|R|L|E|U|S|E|R|L|E|U|G|Y|L|E|U|P|H|E 102  
Db 258 CTTGAGTTCAGAGCTTCAGTGGATATGCAATTCCTTGTTATGGAGAGACTATTTTC 317  
Qy 103 A|L|A|Y|S|P|R|O|M|E|T|A|R|G|S|E|R|Y|S|G|Y|T|Y|V|A|L|T|H|E|T|L|E|U|A|S|P|R|O|P|H|E|Y|S|I|N|I|E|Y 122  
Db 318 GGAAGAAATGCGAGAGAGAGATATATTAACAATCTCGATCTTTTACGACAAATAT 377  
Qy 123 G|Y|L|Y|A|R|G|M|E|T|G|Y|L|E|U|E|U|P|H|E|L|E|P|R|O|A|L|E|U|M|E|G|Y|L|U|M|E|T|P|H|E|T|P|A|I 142  
Db 378 GGCACACGATCGTGGCTTGATGTTCCAGCACTTCTTGTAACATTCCTGACCA 437  
Qy 143 A|L|A|A|I|L|P|H|E|S|E|R|A|L|E|U|G|Y|A|L|A|T|H|I|L|E|S|E|R|V|A|I|L|E|A|S|P|V|A|L|A|S|P|A|A|S 162  
Db 438 GAGCGCATCTTTTCCGACCTGTGTCACACCTGCGTAATCTTGGAAATCGACGAAT 497  
Qy 163 I|L|E|S|E|R|V|A|I|L|E|A|S|P|V|A|L|E|U|I|L|E|A|I|L|E|U|Y|T|H|L|E|U|V|A|G|Y|L|E|U|Y 182

Db 498 GCATCAGTACCCCTGTGCGCTGTATTCGCTATTCACATTCACCGGTGATCTAT 557  
Qy 183 S|E|R|V|A|I|A|L|Y|T|H|A|S|P|V|A|I|G|U|L|E|U|P|H|E|C|Y|S|I|L|P|H|E|I|L|E|U|T|P|I|L|E|S 202  
Db 558 GCAATGCGCTACACTGACGCTGTTCACTATTTTGCATTTTGTGCGTTTGTGGGTTTG 617  
Qy 203 V|A|L|P|R|O|P|H|E|A|L|E|U|S|E|R|I|S|P|R|O|A|L|A|V|A|L|T|H|A|S|P|I|L|E|Y|P|H|E|T|H|A|L|A|H|I|S|A|I 222  
Db 618 GTGCGCGCGCTATATGTCATGATGTGTGGCAAGATATTTCCAGAAATGCA----- 668  
Qy 223 L|Y|S|Y|G|I|N|S|E|R|P|R|O|T|R|E|U|G|Y|T|H|I|L|E|G|U|S|E|R|V|A|L|-----G|U|V|A|I|Y|T|H|T|P|L|E|U 241  
Db 669 -----GGCAGCTCGGATTCGAGAGATTGGAGATTGCAAGAAACATCTCTGTGAT 719  
Qy 242 A|S|P|A|S|P|H|E|L|E|U|L|E|U|M|E|T|L|E|U|G|Y|L|I|P|R|O|T|R|G|I|A|L|A|T|P|H|E|G|I|N|A|R|G|V|A|I 261  
Db 720 GATTGATGCTTCTCTCTTGTCTTTGGAGAAATTCATGGCAAGTGTACTTCCAAAGAGTT 779  
Qy 262 L|E|U|S|E|R|S|E|R|S|E|R|A|T|H|Y|R|A|G|I|N|V|A|L|L|E|U|S|E|R|P|H|E|U|A|L|A|P|H|E|G|Y|C|Y|S 281  
Db 780 CTCTCTCAAAACAGTCTCATGAGACAGAGCTTGTCTGTTGTGGCGGCGCTGGATGC 839  
Qy 282 L|E|U|V|A|L|E|U|P|R|O|A|L|A|I|L|E|C|Y|S|I|L|E|G|Y|A|I|L|E|G|Y|A|L|A|S|E|R|T|H|A|S|P|T|P|A|S|N 301  
Db 840 ATTCTCATGCGCATTCACACAGCGTTGATCGGCAATTGCCAGAAACAGACTGGAGA 899  
Qy 302 G|U|T|H|A|L|A|T|Y|R|G|Y|R|P|R|O|A|S|P|R|O|L|Y|S|T|H|Y|S|G|U|L|U|A|A|----- 316  
Db 900 ATGACTGATTAATTCCTCCATGAAACAAATGGAATGAACTGAATCGATTCACCGGATAG 959  
Qy 317 ---A|S|P|M|E|T|L|E|U|P|R|O|I|L|E|U|G|I|N|Y|L|E|C|Y|S|P|R|O|V|A|I|Y|R|I|L|E|S|E|R|P|H|E 335  
Db 960 AGAAATGATGTGTGTCCTGTTGTGATTCAGTATCTTACCCAAAGATGGGCGCTTAT 1019  
Qy 336 G|Y|L|E|U|G|Y|A|L|A|V|A|L|S|E|R|A|A|A|V|A|L|M|E|T|S|E|R|S|E|R|A|L|A|S|P|S|E|R|I|L|E|U|S|E|R|A|I 355  
Db 1020 GGACTCGGCGCAGTGCCTGCTGTATGATGTCATCGCAGATTCATCTGACTATATGCA 1079  
Qy 356 S|E|R|S|E|R|P|H|E|A|L|A|R|G|A|S|N|I|L|E|Y|R|I|N|L|E|U|S|E|R|P|H|A|R|G|I|N|S|A|L|A|S|E|R|A|P|L|Y 375  
Db 1080 GCATCATATTTGTTCTCAACACCTGGAAGCTCACATTTGCCCTCAGCGGTGAAAAA 1139  
Qy 376 G|U|I|L|E|V|A|T|P|V|A|L|M|E|R|A|G|I|L|E|T|H|V|A|L|E|U|V|A|P|H|E|G|Y|A|L|A|S|E|R|A|L|A|T|H|A|L|A|M|E 395  
Db 1140 GAAATGATATTTGTATGATGATAGCATTCGCTGTGATCATGGCAATCATCATG 1199  
Qy 396 A|L|A|L|E|U|T|H|L|Y|S|T|H|V|A|I|Y|G|Y|L|E|U|T|P|Y|R|L|E|U|S|E|R|S|E|R|A|P|L|E|U|V|A|I|Y|R|I|L 415  
Db 1200 GCACTTACCATTCATCCATCATATGAGGCTTGTGTATCTTTGACAAATTTGGCTACGTC 1259  
Qy 416 I|L|E|L|P|H|E|P|R|O|G|I|L|L|E|U|C|Y|S|V|A|L|E|U|P|H|E|L|E|Y|G|Y|T|H|Y|S|N|H|Y|R|Y|G|Y|A|A 435  
Db 1260 ATACTCTTCTCAACTATATATGTTGTAATATGCCAGTAGCAATACGATGTGCTCA 1319  
Qy 436 V|A|L|A|G|Y|T|Y|R|I|L|P|H|E|G|Y|L|E|U|P|H|E|U|A|G|I|L|E|T|H|G|Y|G|Y|L|U|P|R|O|Y|R|L|E|U|Y 455  
Db 1320 TTGGCGGCTATGCACTGCGTCTTGCTCCGTTGATTTGGAGCGACCACTTGTATCG 1379  
Qy 456 L|E|U|G|I|N|P|R|O|L|E|U|I|L|E|P|H|E|Y|R|P|R|O|G|Y|T|Y|R|S|E|R|A|S|P|L|Y|S|A|N|G|Y|I|L|E|Y|R|A|E|N|G|I 475  
Db 1380 CTGCGACGCTTCTTCATATCAATGAT-----ACGGATGGGGTA-----CAG 1424  
Qy 476 A|R|G|P|H|E|R|O|P|H|E|L|Y|S|T|H|L|E|U|S|E|R|T|V|A|L|T|H|S|E|R|P|H|E|T|H|Y|S|N|I|L|E|C|Y|S|V|A|L|S|E 495  
Db 1425 TATTTCCATTCAGGCACTGATATTTCTTCAATGCTACTATTTCAATGATGATCA 1484  
Qy 496 T|Y|L|E|A|L|A|Y|S|T|H|L|E|U|P|H|E|G|U|S|E|R|G|Y|T|H|L|E|U|P|R|O|L|Y|S|L|E|U|A|S|P|V|A|L|P|H|E|A|S 515  
Db 1485 ATACATTCGGAAGAGTGTTCATTAATCGGACGTTTGTCTCCGAGAGGACGTAAGGAT 1544  
Qy 516 A|L|A|V|A|I|A|-----A|L|A|R|G|H|S|S|E|R|G|U|G|U|A|S|M|E|R|A|S|P|L|Y|S|H|I|L|E|U|V|A|I 532  
Db 1545 TGTGTAGTAATATTCGATAGATCATGTAACCTTCCTCGTCAATGATATGCTGCTGT 1604



QY 533 ArgAnGUaSnllLeuLeu-----AaNGluLeuAlProValLySProArG 548  
Db 1605 AGTAGTGAGACCTTGAAATATGAAAGCTCCAAACGGAACCGGCTCCAGTACATCCGAAC 1664  
QY 549 Gln 549  
Db 1665 CAA 1667  
RESULT 13  
US-10-724-806-7  
; Sequence 7, Application US/10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
; FILE REFERENCE: VBELT:00808  
; CURRENT APPLICATION NUMBER: US/10/724,806  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US/09/911,077A  
; PRIOR FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1985  
; TYPE: DNA  
; ORGANISM: Caenorhabditis elegans  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (9)..(1739)  
; US-10-724-806-7  
Alignment Scores:  
Pred. No.: 4,34e-139 Length: 1985  
Score: 1444.50 Matches: 282  
Percent Similarity: 68.6% Conservative: 93  
Best Local Similarity: 52.0% Mismatches: 145  
Query Match: 48.3% Indels: 31  
Gaps: 9  
US-10-724-806-4 (1-580) x US-10-724-806-7 (1-1985)  
QY 7 GlyLeuValAlAlIleLeuPheTyLeuLeuIlePheLeuValGlyIleTrrAlaIa 26  
Db 24 GGATATCGGCGCATGTTCTTCTTACGCGTCATCTTGTAATATGAGGAGGT 83  
QY 27 TrrLySThrLyAaSerGlyAsnProGluGuaG-----SerGluAlaIe 42  
Db 84 AGAAATCGAAAAGTTCAAAAGAGCTTGATCAGAAAGCCGCGCGCGAGAGAGGTG 143  
QY 43 IleValGlyLyAArgAspIleGlyLeuLeuValGlyIlePheThrMetThrAlaThrTrp 62  
Db 144 ATGTAGCGTGGAGAAACATCGGAAGCTCTTGTCCGAATTTTCAACAATGATCCGACGTG 203  
QY 63 ValGlyGlyLyTyrlIeAsnGlyThrAlaGluAlaValTyrglyProGlyCysGlyLeu 82  
Db 204 GTTGGCGGTGCTTATATCATGGAACCGCGGAGGCTCTGTATATATGA-----GGTCTC 257  
QY 83 AlaTrrAlaIleAlaProIleGlyTySerLeuSerLeuIleLeuGlyLyLeuPhePhe 102  
Db 258 CTGGATGTCAAGCTCCAGTGTGATATGCAATTTCCCTGTATAGGAGAGACTACTTTC 317  
QY 103 AlaLySProMetLeuArgSerLySgLyTyValIThrMetLeuAspProPheLySgIleTy 122  
Db 318 GCAAGAAAGAAATCGGAGAGAGATATATACATGCTCGATCCTTTTACGACAAATTT 377  
QY 123 GlyLySArgMetGlyLyLeuLeuPheIleProAlaLeuMetGlyLyIlePheThrAla 142  
Db 378 GGCCCAACGAATCGGTGGCTGTATGTTCAGACACTTCTTGGAAGCAATTTCTGGACA 437  
QY 143 AlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsn 162

Db 438 GCAGCATTTCTTTGGGCACTTGTGCACACTGTGCTGAATTTCTTGAATGCACATGAT 497  
QY 163 IleSerValIleValSerAlaLeuIleAlaIleLeuTyrrIleValGlyLyLeuTy 182  
Db 498 GCATCAGTACCCCTGTGCGCTGTATCCGATTTCTACATTCACCGGTGGATACAT 557  
QY 183 SerValAlaTyrrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrpIleSer 202  
Db 558 GCAGTCGGGTACACAGACGCTGTTCACTATTTTGCATTTTCGTGGTGGGGTTGC 617  
QY 203 ValProPheAlaLeuSerPheProAlaValIThrAspIleGlyPheThrAlaValHisAla 222  
Db 618 GTGCCGGCGCTATGTGTCATGATGTCGCAAGATATTTCCAGGAATGCA----- 668  
QY 223 LyStryGlnSerProTrpLeuGlyThrIleGluSerVal---GluValTyrrThrTrpLeu 241  
Db 669 -----GGGCACTGATTTGGAGATTGGAGATTGGAAGAAACATCTCTCGATT 719  
QY 242 AspaAsnPheLeuLeuLeuMetLeuGlyLyIleProTrpGlnAlaTyrrPheGlnArgVal 261  
Db 720 GATTGCATGCTTCTCTGCTTGTGGAGAAATTCATGCGCAAGTGTACTTCCAAAGAGTT 779  
QY 262 LeuSerSerSerSerAlaThrTyrrAlaGlnValLeuSerPheLeuAlaIlePheGlyCys 281  
Db 780 CTCTCTCAAAAACATGCTCATGACACACAGACAGCTTGTGTGTGGCGCGCTCGGATC 839  
QY 282 LeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsn 301  
Db 840 ATTTCATGGCGATTCCACCAAGCTTGATCGGTGCAATTCGAGGAACACAGACTGGAGA 899  
QY 302 GlnThrAlaTyrglyTyrrProAspProLySThrLySgIleGluAla----- 316  
Db 900 ATGACTGATTTATTCCTCCATGGAACATGGAACATGAACTGAAGTCAATTCACCGGATAAG 959  
QY 317 ---AspMetIleLeuProIleValLeuGlnTyrrLeuCysProValTyrrIleSerPhePhe 335  
Db 960 AGAAACATAGTGGTCCCGTGGTATTCAGATTCCTTACCCCAAGATGGGTCCCTTTATT 1019  
QY 336 GlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAla 355  
Db 1020 GGACTCGGCGGAGTGTGCGCTGTATATGATCATCTGCAGATTCATCTGACTATACAGA 1079  
QY 356 SerSerMetPheAlaArgAsnIleTyrrGlnLeuSerPheIleArgGlnAsnAlaSerLyS 375  
Db 1080 GCATCAATGTTTGTCTCACAACTCGGAAGCTCAAAATTCGCTCAGCGCTGTAATAA 1139  
QY 376 GluIleValIrrValMetArgIleThrValIleValPheGlyAlaSerAlaThrAlaMet 395  
Db 1140 GAAATGATTAATGTGATGAAATACCATCATCTGTGTGGTATCATGGCAACATCATG 1199  
QY 396 AlaLeuLeuThrLySThrValTyrglyLeuTyrrLyLeuSerSerAspLeuValTyrrIle 415  
Db 1200 GCATTTACCATTCATTCATCATGATGAGCTTGTGTATCTTGTGAGATTTGGTCTACGTC 1259  
QY 416 IleIlePheProGlnLeuLeuCysValLeuPheIleLySgLyThrAsnThrTyrrGlyAla 435  
Db 1260 ATACTCTTCCCTCACTATATGTGTGTATATATGACCAAGTGAAGAAACGATATGCTCA 1319  
QY 436 ValAlaGlyTyrrIlePheGlyLeuPheLeuArgIleThrGlyLyGlyIleProTyrrTy 455  
Db 1320 TTGGCTGTGATGACGTGGTGTCTGTCTGCTGTGATTTGGAAGCGAGCACTTGTATG 1379  
QY 456 LeuGlnProLeuIlePheTyrrProGlyTyrrTyrrSerAspLySAsnGlyIleTyrrAsnGln 475  
Db 1380 CTGGCAGAGCTTCTTCATTTATCATATGAT-----ACGATGGGGTA-----CAG 1424  
QY 476 ArgPheProPheLySThrLeuSerMetValIThrSerPhePheThrAsnIleCysValSer 495  
Db 1425 TATTTCCATTCAGAGCAACGCTATGTATCTTCAATGGCTACTACATTTATATCA 1484  
QY 496 TyrrLeuAlaLyStryLeuPheGluSerGlyThrLeuProLyLeuAspValPheAsp 515



Db 1485 ATACATCGAGAGCTGTTCAATCGGACGTTTGTCTCCGAGTGGACGTAATGGGT 1544  
Qy 516 AlAlaVal-----AlaAghIsSerGIuGluAsnMetAspLysThrIleLeuVal 532  
Db 1545 TGTGTATGTAATATTCGATAGATCATGCCCTTCCGTCAATGTATCGTTGCTGTT 1604  
Qy 533 AGsAsnGIuAsnIleLysLeu-----AsnGIuLeuAlaProValLysProArg 548  
Db 1605 AGTAGAGAACCTTGATATGTAGAGCTCCAAAGCAACCGGCTCCAGTACATCCGAAC 1664  
Qy 549 Gln 549  
Db 1665 CAA 1667  
RESULT 14  
US-11-097-143-41593/c  
; Sequence 41593, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; PRIOR FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 41593  
; LENGTH: 4223  
; TYPE: DNA  
; ORGANISM: DROSOPHILA  
US-11-097-143-41593  
Alignment Scores:  
Pred. No.: 2,57e-135 Length: 4223  
Score: 1413.00 Matches: 305  
Percent Similarity: 59.9% Conservative: 80  
Best Local Similarity: 47.4% Mismatches: 116  
Query Match: 47.2% Indels: 143  
DB: 13 Gaps: 10  
US-10-724-806-4 (1-580) x US-11-097-143-41593 (1-4223)  
Qy 4 HisValGIuGIuLeuValAlaIleIleLeuPheTyLeuLeuIlePheLeuValGIuIle 23  
Db 3217 AATATCGCTGGCGTGGAGACATCGTCTTCTACTCTGATCCGATCGTTGGCAT 3158  
Qy 24 TTPAlaIleATrIlySthrIlyAsnSerGIyAsnProGIuGIuArgSerGIuAlaIleIle 43  
Db 3157 TGGGCGCGTGGCAAG---AAGCAATCCGCGCAATGATTCCGAG-----GAGGAGTCA 3107  
Qy 44 ValGIuGIuArgAspIleGIuLeuValGIuGIuPheThrMetThrAlaTrIlyVal 63  
Db 3106 CTGGCCGAGCGCTCCATCGGCTCTTGTGGGCAATTCACCATGACGGCCACCTGGG 3047

Qy 64 GIuGIuGIuTyrlIeAsnGIuThrAlaGIuAlaValTyrlGIuProGIuCysGIuLeuAla 83  
Db 3046 GGTGGCGGCTATCATCAACGAGCGGCGAGGCTATATAC-----ACATCGGCTTG 2993  
Qy 84 TTPAlaIleAlaProIleGIuTyrlSerLeuSerIle----- 96  
Db 2992 TGTGTCCAGGCTCCATTTGGATACGCTTAAGCTTGTAATTGGTAATCACTTACAG 2933  
Qy 97 -----LeuGIu 98  
Db 2932 ATTCCATTAAAGATTCCTTTTACAAAAAGATTTCTATCTCTCCAACTTAAGT 2873  
Qy 99 GIuLeuPhePheAlaLysProMetArgSerLysGIuTyrlValThrMetLeuAspProPhe 118  
Db 2872 GGCATCTTCTTGGCCATCCATCCGATCCGAGGAGGTATCATCCATGTGGATCCGTTG 2813  
Qy 119 LysGlnIleTyrlGIuArgMetGIuGIuLeuPheIleProAlaLeuMetGIuGIu 138  
Db 2812 CAGGATTCCTTGGTGGAGGAGGATGCTCTTCTGCGCCCTCTATGCGGTGAG 2753  
Qy 139 MetPheTrpAlaAlaIlePheSerAlaLeuGIuAlaThrIleSerValIleIleAsp 158  
Db 2752 GTCTTTGGGACCGCGCATCTCGCTGCACCTTGGCGCATCTATCGGTATCATCGAC 2693  
Qy 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrlThrLeuVal 178  
Db 2692 ATGATCACCGGACCTCGGTGATCTGCTCCTCGATCGGCATCTTCACACATGTTG 2633  
Qy 179 GIuGIuLeuTyrlSerValAlaTyrlThrAspValAlaGIuLeuPheCysIlePheIleGIu 198  
Db 2632 GGTGGACTGATCTCCGTGGATGATGAGAGCTGTACAGTGTTCGTGCATCTTCATCGGT 2573  
Qy 199 LeuTrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGIuPheThr 218  
Db 2572 CTGTGATGTGCATCTCTTCCCTGCGCAAGACACATCGGCGACCTG----- 2522  
Qy 219 AlAlaValHisAlaLysTyrlSerProTrpLeuGIuThrIleGIuSerValGIuAlaTyrl 238  
Db 2521 -----AGTACCTGGAGGTGATGATTGGGACAGCTGAGCTTAAGAAAGCATTTGG 2471  
Qy 239 ThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGIuGIuLeuProTrpGlnAlaTyrlPhe 258  
Db 2470 CTGTACATGACATACCGCTTGTGCTGCTGTTGGTGACATTCCTGGCGGTCTACTTC 2411  
Qy 259 GlnArgValLeuSerSerSerSerAlaThrTyrlAlaGIuValLeuSerPheLeuAlaAla 278  
Db 2410 CAGCGGAGTGTGTCCAGCAAAACGGCAGGAAGGCGCAAGCTTGTCTTATGTTGCGACC 2351  
Qy 279 PheGIuCysLeuValMetAlaLeuProAlaIleCysIleGIuAlaIleGIuAla----- 296  
Db 2350 GCCGATGATTTGTATGGCATTTCCCGGTGCTCATCGAGCATGTCGAAGGCTACA 2291  
Qy 296 ----- 296  
Db 2290 CGTAGTCTTAGTACGACGACACTAGCAAGTAATAACAGATATCCCAAAATAGAAAGC 2231  
Qy 296 ----- 296  
Db 2230 ACATATCTTGTGGTGGGGGAGATTAATCTATGACCTTGTGACTATGCTATTTTCTT 2171  
Qy 297 ---Ser---ThrAspTrpAsnGIuThrAlaTyrlGIuTyrlProAspProLysThrLysGIuGIu 315  
Db 2170 TTCTGTACAGCTTGGAGACAGACATTTACAGAGGACCTTATCCCTGACCTGGCGAGA 2111  
Qy 315 uAlaAspMetIleLeuProIleValLeuGIuGIuTyrlCysProValTyrlIleSerPhePhe 335  
Db 2110 GACGACATGATTTCTGCGATGTGTGCGAGTACCTTCAAGCTTCACTTCTGATCTTCTT 2051  
Qy 335 eGIuLeuGIuAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAl 355  
Db 2050 TGGATTGGGGCGCTTTCCGCGCGCGGTGATGTCTCCGCGGACATCTCCGTTGCTTCGCG 1991  
Qy 355 aSerSerMetPheAlaArgAsnIleTyrlGIuLeuSerPheArgGlnAsnAlaSerAspLy 375

```
Db      1990  CGCCCTCCAGTTCGTCGCGAAGCTACAAATGATTTTCCCTCGAAGGCGCTCCGAGAT 1931
Oy      375  sgluileValTrrPValmetArgIleThrValleuValPheglYalaSerAlaThAlaMe 395
Db      1930  GGAATTCATTGGGTGATGCGAGTCCGATCATTTGTGGGCAATCCGGTAACATCAT 1871
Oy      395  lAlaleuLeuThrIlysrhValTyrGlyLeuTrpTyr----- 407
Db      1870  GGCCCTCACCAATTCCTCCATCTACGATTGTGGTGG- CGTAGAAGAGATCATTTATTCACC 1812
Oy      408  -----LeuSe 409
Db      1811  CTCCTCAGCTCATTTTGTCTAAGTCCCTGCTCTTCTTATCTTTCCGAGAGCTGATGG 1752
Oy      409  rSerAspLeuValTyrIleIleIlePheProGlnleuLeuCysValleu---PheIlely 428
Db      1751  CTCGGATCTGGTCTACGTCATCTCTTCCCGACGCTACTGATGTGTGTCACCTTCAAGAA 1692
Oy      428  sglYThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheglYleuPheLeuArgIleTh 448
Db      1691  GCACGCAACACAGTACGCGAGCGCTGTCGACATACATTGGCCCTGGCCATCCGACTGTC 1632
Oy      448  rglYglYgluProTyrTyrLeuGlnProleuIlePheTyrProGlyTyrTyrSerAs 468
Db      1631  GGGCGGTAGGCGCATCTTGGGAGCTGCTCCATTCATTCAGATCCCGGCTACGACGAGA 1572
Oy      468  pLysAsnGlyIleTyrAsnGlnArgPheProPheIlysrhLeuSerMetValThrSerPh 488
Db      1571  GACCAAG-----GAGCANATGTTCCTTCCGACATGAGGCATCTCTCAAGCT 1521
Oy      488  ePheThrAsnIleCysValSerTyrLeuAlaIys----- 499
Db      1520  GGTACGCTCATCTGCTCTCTCTGAGCTAAGATAGATTAAGCTATGTAATTACAT 1461
Oy      499  ----- 499
Db      1460  AGTACTACATTTGATTATATTATGTAATTAAATTAATTAATTAATTAATTAATTAAT 1401
Oy      500  -----TyrLeuPhegluSerGlyThrleuProProlYsLeuAspValPheAspAl 516
Db      1400  CAATTTCAGAAATGATGTTGATGTCGCGCAAGTTGCGCCGACGCTACGACTACTTCCGCTG 1341
Oy      516  aValVal 518
Db      1340  TGTGCTC 1334

RESULT 15
US-11-097-143-41594
; Sequence 41594, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41594
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-41594

Alignment Scores:
Pred. No.: 4.55e-135 Length: 1729
Score: 1405.00 Matches: 286
Percent Similarity: 69.1% Conservative: 71
Best Local Similarity: 55.3% Mismatches: 105
Query Match: 46.9% Indels: 55
DB: 13 Gaps: 8

US-10-724-806-4 (1-580) x US-11-097-143-41594 (1-1729)
Oy      4  HisValGluGlyLeuValAlaIleIleleuPheTyrLeuLeuIlePheLeuValGlyIle 23
Db      7  AATATCGCTGGCGGTGTAGCATGCTGCTCTTCTACCTCTGATCTGCTGCTTGGCATT 66
Oy      24  TrpAlaAlaTrpLysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle 43
Db      67  TGGCGCGGTTCGCAAG--AAGCAGTCCGCAATGATTCGAG--GAGGAGGTCATG 117
Oy      44  ValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpVal 63
Db      118  CTGGCGGAGCGCTTCATGGCTCTTCTGAGCATCTTCAACATACAGCCACTGGGTG 177
Oy      64  GlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAla 83
Db      178  GGTGGCGGCTACATCAAGCGCAGCGAGGCTATATAC-----ACATCGGCTGCTGTC 231
Oy      84  TrpAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyLeuPhePheAla 103
Db      232  TGGTCCAGGCTCCCATTTGGATACCGCTTAAGCTTGGATTGGTGGCATCTTCTTGGCC 291
Oy      104  LysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGly 123
Db      292  AATCCCATGCGCAGACAGGATTACATCAACAGTTGATCCGTTGACAGGATTCCTTTGGT 351
Oy      124  LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPheTrpAlaAla 143
Db      352  GAGCGGATGGAGAGATTTGCTCTCTGCGCGCTCATAGCGGTGAGGTCTTTTGGCAGCC 411
Oy      144  AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle 163
Db      412  GGCATCTGCTGCTGCACTTGGCGCCACTATCGGTGATCATCGATGATGATACCGCACCC 471
Oy      164  SerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyLeuTyrSer 183
Db      472  TCGGTGATCTGCTCTCTCTCGATCGCATCTTCTTCAACACTTGTGGGTGACTGATCC 531
Oy      184  ValAlaTyrThrAspValAlaGlnleuPheCysIlePheIleGlyLeuTrpIleSerVal 203
Db      532  GTGGCGTATAGGAGCATGATTCAGTTGTTGCACTTCATCGCTGAGTGTGACATT 591
Oy      204  ProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaIys 223
Db      592  CCTTTCGCTGAGC----- 606
Oy      224  TyrGln-SerProTrpLeuGlyThrIleGluSerValGluValTyrThrTrpLeuAspAs 243
Db      607  ---AACGAGCAGCTGGGACGCTGAGTG-----ACCTGG----- 637
Oy      243  nPheLeuLeuLeuMetLeuGlyIleProTrpGlnAlaTyrPheGlnArgValleuSe 263
Db      637  ----- 637
```



***This Page Blank (uspio)***

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 10, 2006, 14:48:06 ; Search time 404 Seconds  
(without alignments)  
2799.791 Million cell updates/sec

Title: US-10-724-806-4  
Perfect score: 2993  
Sequence: 1 MFPHVGLVAILFLYFLPL.....EALLDVSSPEGSGETEDNLQ 580

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues  
Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+.p2n.model -DRV=xlh  
-Q=/abs/ABSSWEB.spool/US10724806/runatc.10072006.140301.7914/app\_query.fasta.1  
-DB=PubMed\_Applications\_NA\_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -HOST=abs807  
-USER=US10724806@CGN1.1.137@runatc.10072006.140301.7914 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA\_New:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/BCT\_NEW\_PUB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the local score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	296	9.9	2449	US-11-266-748A-28918	Sequence 28918, A
2	296	9.9	2449	US-11-266-748A-56547	Sequence 56547, A
3	282.5	9.4	2273	US-11-266-748A-29027	Sequence 29027, A
4	257	8.6	20817	US-11-021-837-16	Sequence 16, Appl
5	255	8.5	15455	US-10-471-571A-1417	Sequence 1417, Ap
6	245.5	8.2	3227	US-11-266-748A-56747	Sequence 56747, A
7	235	7.9	95963	US-11-021-837-39	Sequence 39, Appl
8	206.5	6.9	2211	US-11-217-529-1082	Sequence 1082, Ap

9	205.5	6.9	2208	US-11-217-529-76899	Sequence 76899, A
10	197.5	6.6	1539	US-10-471-571A-1991	Sequence 1991, Ap
11	191	6.4	2984	US-11-266-748A-26072	Sequence 26072, A
12	184	6.1	38840	US-11-021-837-27	Sequence 27, Appl
13	169	5.6	3844	US-11-266-748A-27661	Sequence 27661, A
14	168	5.6	1578	US-11-266-748A-187885	Sequence 187885, A
15	143.5	4.8	936	US-11-266-748A-21747	Sequence 16097, A
16	143.5	4.8	936	US-11-266-748A-359132	Sequence 359132, A
17	143	4.8	933	US-11-266-748A-386971	Sequence 386971, A
18	143	4.8	933	US-11-266-748A-442511	Sequence 442511, A
19	143	4.8	933	US-11-266-748A-442511	Sequence 442511, A
20	142.5	4.8	2171	US-11-218-305-9650	Sequence 9650, Ap
21	137	4.6	1434	US-10-471-571A-3889	Sequence 3889, Ap
22	136	4.5	2513	US-10-449-902-20129	Sequence 20129, A
23	133.5	4.5	1966	US-11-266-748A-27363	Sequence 27363, A
24	131	4.4	59143	US-11-021-837-34	Sequence 34, Appl
25	129.5	4.3	2429	US-10-449-902-23815	Sequence 23815, A
26	127	4.2	1000	US-11-266-748A-291922	Sequence 291922, A
27	127	4.2	1000	US-11-266-748A-343351	Sequence 343351, A
28	127	4.2	1000	US-11-266-748A-403483	Sequence 403483, A
29	127	4.2	1000	US-11-266-748A-474529	Sequence 474529, A
30	124	4.1	1209	US-10-471-571A-2713	Sequence 2713, Ap
31	123	4.1	1730	US-11-174-3078-5197	Sequence 5197, Ap
32	122.5	4.1	6570	US-11-021-837-10	Sequence 10, Appl
33	121	4.0	1928	US-10-449-902-20061	Sequence 20061, A
34	121	4.0	3070	US-11-218-305-1841	Sequence 1841, Ap
35	120	4.0	1377	US-10-471-571A-471	Sequence 471, Ap
36	119.5	4.0	2614	US-11-216-545-8626	Sequence 8626, Ap
37	119	4.0	2186	US-11-266-748A-32772	Sequence 32772, A
38	119	4.0	2237	US-11-266-748A-281298	Sequence 281298, A
39	118.5	4.0	1000	US-11-266-748A-288126	Sequence 288126, A
40	118.5	4.0	1000	US-10-449-902-239555	Sequence 239555, A
41	117.5	3.9	2216	US-10-449-902-24763	Sequence 24763, A
42	117	3.9	1592	US-11-216-545-3055	Sequence 3055, Ap
43	116.5	3.9	1941	US-11-174-3078-1767	Sequence 1767, Ap
44	116.5	3.9	100445	US-11-021-837-40	Sequence 40, Appl
45	116	3.9	1470	US-11-217-529-396	Sequence 396, Appl

#### ALIGNMENTS

RESULT 1  
US-11-266-748A-28918  
; Sequence 28918, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266, 748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: BP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662, 276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700, 293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 28918

```

; LENGTH: 2449
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-28918

```

## Alignment Scores:

```

Pred. No.: 3.45e-22 Length: 2449
Score: 296.00 Matches: 148
Percent Similarity: 38.1% Conservative: 99
Best Local Similarity: 22.8% Mismatches: 222
Query Match: 9.9% Indels: 180
Gaps: 29

```

US-10-724-806-4 (1-580) x US-11-266-748A-28918 (1-2449)

```

OY 11 IletleuPheTyrlleuLeuIlePheLeuValGlyIleTPrAlaIleTPrAlaIleThyrls 30
DB 104 ATCGTTATCTACTTGGTGTAGTATGCGCCGTGAGCTGGCTATGTTTCCACC--- 160
OY 31 AenserGlyAsnProGluGluArgSerGluAlaIleIleValGlyIleYArgAspIleGly 50
DB 161 AATCGTGGGACT---GTTGAGGCTTCTTCCTGGCAGGCCGAGATGTGTG 208
OY 51 LeuLeuValGlyIlePheTherThraIleThrPValGlyIleYTrIleAsnGly 70
DB 209 TGGTGGCCGATTTGGAGCCCTCCTTTGTCTAGTAACTTGGAGTGGCCACTTTGTGGGG 268
OY 71 ThrAlaGluAlaValTyrlGlyProGlyCysGlyLeuAlaTPrAlaIleHisAlaProIleGly 90
DB 269 CTGGCC---GGAGCTGGGGAGCTTCAGCATGCCCATTTGGAGGC 310
OY 91 TyrSer-----LeuSerLeuIleLeuGlyIleuPhePheAlaIleAlaIle 105
DB 311 TTGTGATGGAATGCCCTGGTTTGGTGTGTGTGCTGGCTGGCTGTTGTC-----CCC 364
OY 106 MetArgSerIys---GlyTyrlValTherMetLeuAsnProPheIleGlnIleTyrlGlyIys 124
DB 365 ATCTATATTAAGGCTGGGGGTGTGTGAACATGCCAGATACCTGAGG-----AAG 412
OY 125 ArgMetGlyIle-----LeuLeuPheIle 132
DB 413 CGGTTTGGAGGCGACGCGATCCAGGTACCTTCCCTTCGTCCCTGGCTGTCTTACATT 472
OY 133 ProAlaLeuMetGlyIleuPheThrPAlaIleAlaIlePheSerAlaLeuGlyAlaThr 152
DB 473 TTCACCAAGATCTCGGACAGACATCTTCGCGGGGCCATATTC----- 514
OY 153 IleserValIleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAla 172
DB 515 ATCAATCTGGCCTTGAAGCCTGAATCTGTATTTAGCCATCTTTCCTTATTGGCAATCACT 574
OY 173 IleLeuTyrlleuLeuValGlyIleuTyrlSerValAlaTyrlThraSpValAlaGlnLeu 192
DB 575 GCCCTTTCACAAATTAACAGGGGGCTGGGGCGGTGATTTACACGAGACACTTGGACAG 634
OY 193 PheCysIlePheIleGlyLeuTyrlIleSerValProPheAlaLeuSerHisProAlaVal 212
DB 635 GTGATCATGTGGTGGGTCTTAACTCTGACTGGGTTGCTTTTACAGAACTGGGA--- 691
OY 213 ThrAspIleGlyPheThraIleValHisAlaIleTyrlGlnSerProTyrlleuGlyThrIle 232
DB 692 -----GGCTATGAGCCCTTTCATGGAAAGTTCATGAAAGCC---ATTCCAACTATA 739
OY 233 GluSer-----ValGluValTyrlThr---TrpLeuAspAsnPhe 244
DB 740 GTGTCTGTAGGCAACACACCTTTCAGGAAATAATGCTACATCCAAAGGCCGACCTCTTC 799
OY 245 LeuLeuLeu-----MetLeuGlyIleTyrlPheTrp----- 254
DB 800 CACATCTTCCAGATCCCTTCACGGGAGACCTCCCATGGCTCGGTTCATCTTTGGGATG 859
OY 255 -----GlnAlaTyrlPheGlnAlaTyrlValLeuSer 263

```

```

DB 860 TCATCTTACCTTGTGTGTAAGTGCACAGATAGATCTTGTGCAGCCCTGCTTCA 919
OY 264 SerSerSerAlaThrTyrlAlaGln-----ValLeuSerPheLeuAlaIlePhe 279
DB 920 GCCAAGATATGTCTCACGTGAAGGGTGGCTGCATCTGTGGGTATCTAAAGCTGATG 979
OY 280 GlyCysLeuValMetAlaLeuProAlaIle----- 289
DB 980 CCCATGTTTCATCATGTGATGTCAGGAATATGACGCCCATTTCTGTACACAGAAAAATT 1039
OY 290 -----CysIleGlyAlaIleGlyAlaSerThraSpTrpAsnGln 302
DB 1040 GCCTGTGCTGCTTCAGATGTGAGAAATATTTCCGTACCAAGCTTGCTATCCAA 1099
OY 303 ThrAlaTyrlGlyTyrlProAspProIysThyrlGluGluAlaAspMetIleuProIle 322
DB 1100 ATCGCCTAT-----CCAACC 1114
OY 323 ValLeuGlnTyrlleuCysProValTyrlIleSerPhePheGlyLeuGlyAlaValSerIle 342
DB 1115 TTAGTGTGAGACTTCATGCGCAATGAGCTGAGGCTGATGCTATCATGTGTGCTC 1174
OY 343 AlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsn 362
DB 1175 TCCCTCATGAGCTCCCTTCACTTCACTTCAACAGCCGACACCCCTTTCACCATGGAC 1234
OY 363 IletTyrlGlnLeuSerPheArgGlnAsnAlaSerAspIysGluIleValTyrlValMetArg 382
DB 1235 ATCTAC---GCCAAGGTCCGCAAGAGAGCATCTGAGAAAGACTCATGATTCGCGGAAG 1291
OY 383 IletThrValLeuVal---PheGlyAlaSerAlaThrAlaMetAlaLeuLeuThyrlsThr 401
DB 1292 TTGTTTATCTCGTGCTGATGATTCGATGCAATGCCCTGGGCCCATTTGTGATGCACGA 1351
OY 402 ValTyrlGly-----LeuTyrlTyrlleuSerSerAspLeuValTyrlle-----IleIle 417
DB 1352 CAAAGTGGGCAACTCTTGATTTACATTCACATCCATCCATGATCTTGGACCAACCATTT 1411
OY 418 PheProGlnLeuLeuCysValLeuPheIleIysGlyThraAsnThyrlGlyAlaValAla 437
DB 1412 GCGGCTGTCTTCCTGCTGCTGATTTTCTGAAAGAGCATGATGAGCCAGAGCCCTTTGG 1471
OY 438 GlyTyrlIlePheGlyLeuPheLeuArgIle----- 447
DB 1472 GGACTGATCTTGAAGCTTCTGATTTGGATTTTCACGTATGATTAAGTTCAGTTGCTTATGGA 1531
OY 448 ThrGly-----GlyGluProTyrlleu 454
DB 1532 ACCGGAGCTGCATGAGACCCAGCAACTGTCCACGATTAATCTGTGGGTGCACACTTGG 1591
OY 455 TyrlleuGlnProLeuIlePheTyrlProGlyTyrlTyrlSerAspIysAsnGlyIleTyrlAsn 474
DB 1592 TACTTGGCATTATTCCTCTTC----- 1612
OY 475 GlnArgPheProPheIleThraSerMetValThraSerPhePheThraAsnIleCysVal 494
DB 1613 -----GCCATTTCTTTCATCCATGCTGTGTCTATC 1642
OY 495 SerTyrlleuAlaIleTyrlleuPheGluSerGlyThraLeuProIleTyrlleuAspValPhe 514
DB 1643 TCCCTCTCCACCAACCATT-----CCGAGTGTGATTTCTTAC 1681
OY 515 AspAlaVal-----ValAlaArgHisSerGluGluAsnMetAspIysThraIleLeuVal 532
DB 1682 CGTGTGTGTGAGGCTGCGCAACAGCAAAAGAGGCTGATTTGAC-----CTGGATGCG 1735
OY 533 ArgAsnGluAsnIleIleLeuAsnGluLeuAlaProValIysProArgGlnSerIleuThr 552
DB 1736 GAAAGAGGAACATCCAGAAAGC-----CTTAAAGAGACATTGAA 1777
OY 553 LeuSerSerThraPheThraAsnGlu 561
DB 1778 ATAGAAACACAAGTTCTCGAGAGAAA 1804

```



Db 1412 GGGGCTGCTTCCTGCTGCTATTTTCTGGAAGAGTCATAGCCAGAGCCTTTGG 1471  
QY 438 G|Y|T|Y|L|E|P|H|E|G|L|Y|L|E|U|P|H|E|U|A|R|G|L|E|----- 447  
Db 1472 GGACTGATCTCTAGGACTTCTGTATTTGGAATTTACCGATGATTTACTGAGTTTGTATGGA 1531  
QY 448 ThrGly-----GlyGluProTyrLeu 454  
Db 1532 ACCGGAGCTGCATGATGAGCCAGCAACTGTCCACGATTATCTGGGGTGCACCTATGG 1591  
QY 455 TyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspLysAsnIlyIleTyrAsn 474  
Db 1592 TACTTGGCATTAATCTCTTC----- 1612  
QY 475 GlnArgPheProPheLysThrLeuSerMetValThrSerPheThrAsnIleCysVal 494  
Db 1613 -----GCCATTTCTTTTCATCACCATCGTGGTCATC 1642  
QY 495 SerTyrLeuAlaIleTyrTyrLeuPheGluSerGlyThrLeuProProlLysLeuAspValPhe 514  
Db 1643 TCCCTCCCTCACCAACCCTAT-----CCGAGTGTGATCTCTAC 1681  
QY 515 AspAlaVal-----ValAlaArgHisSerGluGluAsnMetAspLysThrIleLeuVal 532  
Db 1682 CGTCTGTGTGGAGCTGCGCAACAGCAAGAGGAGCGTATGAC-----CTGATGCG 1735  
QY 533 ArgAsnGluAsnIleLysLeuAsnGluLeuAlaProValLysProArgGlnSerLeuThr 552  
Db 1736 GAAGAGGAAACATCCAGAAAGGC-----CTTAAGAGACCATTTGA 1777  
QY 553 LeuSerSerThrPheThrAsnLysGlu 561  
Db 1778 ATGAAACACAAAGTTCTGAGAAAGAA 1804

RESULT 3  
US-11-266-748A-29027  
; Sequence 29027, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcription Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 29027  
; LENGTH: 2273  
; TYPR: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-29027

Alignment Scores: 9.31e-21 Length: 2273

Score: 282.50 Matches: 143  
Percent Similarity: 36.2% Conservative: 92  
Best Local Similarity: 22.0% Mismatches: 240  
Query Match: 9.4% Indels: 175  
DB: 8 Gaps: 23

US-10-724-806-4 (1-580) x US-11-266-748A-29027 (1-2273)

QY 8 LeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyIleTyrAlaIleTyr 27  
Db 96 ATCTAGCATGCTGCATATTCTCTGCTGGCATGAGCTGGCTGTGGCTGATGTC 155  
QY 28 LysThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyIleTyr 47  
Db 156 AGAACCC--AACAGAGCACT-----GTGGCGGCTACTTCTGGCAGAGCGC 200  
QY 48 AspIleGlyLeuLeuValGlyPheThrMetThrAlaThrPvalGlyIleTyr 67  
Db 201 AGCATGTGTGTGGCGCGTGGGCTCTCTTCCGACAGACATGCGAGTGGCCAC 260  
QY 68 IleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTyrAlaHisAla 87  
Db 261 TTTGTGGCGCTGGCAGGAGCT-----GGCGTGCAGTGGCTGTGGCTGGATTC 314  
QY 88 ProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPheAlaLysPromeArg 107  
Db 315 GAGTGAATGCGCTCTGCTGTGCTGCTGAGCGCTGCTGTTGACCCGCTGACTG 374  
QY 108 SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGlyLysArgMetGly 127  
Db 375 ACAGGGGGGTCATACCATG-----CCACGATACCTGGCAAGGCTTGGCG 422  
QY 128 Gly-----LeuLeuPheIleProAlaLeu 135  
Db 423 GGCGCGCGCATCCGCTCTACCTGTGCTCTCTCTTCTCTGTACATCTCACCAG 482  
QY 136 MetGlyGluMetPheTyrAlaIleAlaIlePhe-----SerAlaLeuGlyAlaThrIle 153  
Db 483 ATCTAGTGACATGTTCTCCGAGCTGTATTATTCACAGAGGCTGTGGGTAACATC 542  
QY 154 SerValIleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAlaIle 173  
Db 543 TATGCC-----TCCGTCATCGCGCTTGGGATCACCATG 578  
QY 174 LeuTyrThrLeuValGlyGlyLeuTyrSerValAlaTyrThrAspValValGlnLeuPhe 193  
Db 579 ATTACAGGTGACAGGAGGCTGGCGCTGATGTACAGGACACGCTACAGACTTC 638  
QY 194 CysIlePheIleGlyLeuTyrIleSerValProPheAlaLeuSerHisProAlaValThr 213  
Db 639 GTCAATTCGGGGGGGCGCTGCATCTCATGGGTTAGCGCTTCCACGAGGTGGC----- 692  
QY 214 AspIleGlyPheThrAlaValHisAlaLysTyr----- 224  
Db 693 -----GGTATTCGGGCTCTTCCACAAATACCTGGGAGCAGCGACTTCGTGACGGTG 746  
QY 225 ---GlnSerProTyrLeuGlyThrIleGluSerValGluValTyrThrTyrLeuAspAsn 243  
Db 747 TCCGAGATCCAGCGCTGGAAACATCTCCAGCTTCTGCTATGACCCCGGCCGATCC 806  
QY 244 PheLeuLeuLys-----MetLeuGlyGlyIleProTyr----- 254  
Db 807 TACACCTGTCTCCGACCCCGGTGACCGGGAGATGTGCGGTGGCCCGGCTGCTCGGA 866  
QY 255 -----GlnAlaTyrPheGlnArgValLeu 262  
Db 867 CTCACATGCTGTGGGCTGTGACTGTGACGACGACGACGATGTGTCAAGGCTGCTG 926  
QY 263 SerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePhe----- 279  
Db 927 GCCGGGAAGAGCTGACCCACATCAAGGCGGCTGATCTGTGTGGTACTGTAAGCTG 986  
QY 279 ----- 279



```

Db      987 AGCCCATGTTCTCATGTCATGCGAGCATGATCAGCGCATCTGTATCCAGACGAG 1046
Qy      280 ---GlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThr 298
Db      1047 GTGGCGTGGGTGGT-----CCTGAGGTGTGAGGCGCGTGTGCGGACGAGGTG 1097
Qy      299 AspTrpAsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGluGluAlaAspMet 318
Db      1098 GGCTGCTCCAAATCGCCTAC----- 1118
Qy      339 IleLeuProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGly 338
Db      1119 -----CCGCGGTCTGCTGGAAGCTCATGCCCACGCTGCGGACGACTCATGCTGGCG 1172
Qy      339 AlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerMet 358
Db      1173 GTATGCTGCGCGCTCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
Qy      359 PheAlaArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleVal 378
Db      1233 TTCACCATGAGACTTAC---ACGCGCTGCGGCGACGCGCGGCGACCGCGAGCTGCTG 1289
Qy      379 TrpValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeu 398
Db      1290 CTGGTGGAGCGGCTGTGGGTGTTCATGCTGTAGTGTGCGGTGCGCTTCCGCTTCCGCTG 1349
Qy      399 ThrLysThrValTyrGly-----LeuTyrTyrLeuSerSerAspLeuValTyrIle 415
Db      1350 GTGCAGCGCGACGAGCGCGGCGACTTTCGATTATCATCGACGACTCTTACCTACCTG 1409
Qy      416 Ile-----IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTyr 433
Db      1410 GCACCGCCCGTGTCCCGCTTCTGCTGCTGCGGCTTCTGCTGCGGCGGCTTATAGACGAG 1469
Qy      434 GlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly----- 449
Db      1470 GGCGCTTCTGGGGACTCATCGGGGCGCTGCTGATGAGCGCTGCGACGCTGATCCCGAG 1529
Qy      450 -----Gly 450
Db      1530 TTCTCTTGGCTCGGCGAGCTGTGTGACGCCCTGCGGCTGCGGCTTCTCTGCTGCGG 1589
Qy      451 GluProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspLysAsn 470
Db      1590 GTGCACCTACTCTACTCCCATGCTGCTGCTTCTC----- 1625
Qy      471 GlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThr 490
Db      1626 -----TGCTGTGGCTCTCTCAACCCCTCACGAGTC 1652
Qy      491 AsnIleCysValSer-----TyrLeuAlaLysTyrIlePheGluSerGly 505
Db      1653 TCCCTGTGACCGCGCCATCCCGAGAAAGACACTTCACGCGCTGTGCTTCACTCCGG 1712
Qy      506 ThrLeuProProLysLeuAspValPheAspAlaValAlaLysHisSerGluGluAsn 525
Db      1713 CATAGAGAGAGAGAGAGAGAGAGCTGATGCT-----GATGAGCAG 1754
Qy      526 MetAspLysThrIleLeuValArgAsn-----GluAsnIleLysLeuAsnGlu 541
Db      1755 CAAGGCTCTCACTCCTGTACAGATGGGTGCCAGAGAGTCCATGAGATGATGAG 1814
Qy      542 LeuAlaProValLysProArgLysSerLeu 551
Db      1815 -----CCCGAGGCGCGGACCAAGCCTC 1838

```

```

; APPLICANT: Manning, Paul
; APPLICANT: McLaughlin, Robert
; APPLICANT: McCormack, Kathleen
; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
; FILE REFERENCE: 101324-US
; CURRENT APPLICATION NUMBER: US/11/021,837
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/533534
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/600660
; PRIOR FILING DATE: 2003-08-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 16
; LENGTH: 20817
; TYPE: DNA
; ORGANISM: Staphylococcus saprophyticus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15033)..(15033)
; OTHER INFORMATION: n is a, c, g, or t
US-11-021-837-16

Alignment Scores:
Pred. No.: 1.8e-16 Length: 20817
Score: 257.00 Matches: 120
Percent Similarity: 42.4% Conservative: 109
Best Local Similarity: 22.2% Mismatches: 215
Query Match: 8.6% Indels: 96
DB: Gaps: 23

US-10-724-806-4 (1-580) x US-11-021-837-16 (1-20817)
Qy      9 ValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyTyrLeuAlaIleTyrLys 28
Db      18837 ATCAGCTGTGGTGTATTTTGTGGTGTGCTAGTATGGGACTATGAGCTTT----- 18890
Qy      29 ThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyArgAsp 48
Db      18891 AATCATCTACCGGTATACCTCAGTGA-----TATATGCTGGGCGGTATAGAAC 18938
Qy      49 IleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTyrValGlyGlyTyrIle 68
Db      18939 ATGGTCATATGATTAATCATTAATTCGACGCGCATCGACATGATGATGATGAT 18998
Qy      69 AsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaThrAlaHisAlaPro 88
Db      18999 ATGGGATTACCTGTGATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 19052
Qy      89 IleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPhePheAlaLysProMetArg--- 107
Db      19053 ATTTGATTAACTTGGCGCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19112
Qy      108 -----SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGly 123
Db      19113 TATACAGAAAGAGCGGAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19172
Qy      124 LysArgMetCysGlyLeuLeuPheIleProAlaLeuMetCysGluMetPheThrAla 143
Db      19173 GATTAATCTTAATTCGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19232
Qy      144 AlaIlePheSerAlaLeu-----GlyAlaThrIleSerValIleIleAspValAsp 160
Db      19233 TATACGATTTCTGGAATGTATACAGCGGTATATTAATTAATTAATTAATTAATTAAT 19292
Qy      161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180
Db      19293 TATCAATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19352
Qy      181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200
Db      19353 TATTTAGCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 19409

```

```

RESULT 4
US-11-021-837-16
; Sequence 16, Application US/11021837
; Publication No. US20060140972A1
; GENERAL INFORMATION:
; APPLICANT: Alm, Richard

```



```
QY 239 -----ThrTrpLeuaspasnPheLeuLeuMetLeuGlyGlyIle----- 252
DB 709 AAACCTACAAATTTAATTTAATTAAGGTTATATCATTTATAGATTTATCTCTATTT 768
QY 253 ProTrp--GlnAlaTyrrPheGlnArg-----ValLeuSerSerSerSerAlaThr 268
DB 769 TCATGGGGATTGATGTTATTTGGTCAACCTCATATCATTTGTAAGTTTATGCTTAA 828
QY 269 TyrrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAla----- 285
DB 829 TCACACAAATGCTACTTAAGCTAGACCTTTAGATTTAGCTGATGCTGTTGTTA 888
QY 286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTy 305
DB 889 TTAGGGCGCTGGCGCTTGTGTTTAAACGATTTGCACTTC-----GTACCTGCTTAT 939
QY 306 -----GlyTyrrProaspProlysthrLysGluGlnAlaAspMetIleLeuProIleVal 323
DB 940 CATATTAACTAGAAAGATCTCTGAGACA-----TTATTCACTCTGATG 981
QY 324 LeuGlnTyrrLeuCysProValTyrrIleSerPhePheGlyLeuGlyAlaValSerAlaAla 343
DB 982 AGTCAAGTACTCTTCATCTCTTGAAGTGTGTTTACTTCTGCTGCAATCTAGCTGA 1041
QY 344 ValMetSerSerAlaaspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363
DB 1042 ATTATGAGACACATTTCTTCACAAATTAATTGTAACCTAGTCACTAACGAGACCTT 1101
QY 364 TyrrGlnLeu-----SerPheArgGlnAsnAlaSerAspLysGluIleValTyrrVal 380
DB 1102 TATTAATTAATTCTGTGTGTAAGAAAGAAAGCTAAACCCACCAAAAGAAATTTGTATGATT 1161
QY 381 MetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLys 400
DB 1162 GGAAGATTATCTGTATTAAGTGTAGCAATTTGTCATCGCGATGCA----- 1209
QY 401 ThrValTyrrGlyLeuTyrrTyrrLeuSerSerAspLeuValTyrrIleIle----- 416
DB 1210 -----TGGAAATCCAAACGACACAACTTAACCTTAAGTAAAGCTTGG 1254
QY 417 -----IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThr 430
DB 1255 GCCGGATTGGTGATCGTTCACTAGTCCACTTGTGCTATTGCACTTAAGGAAAGCTTGG 1314
QY 431 AsnThrTyrrGlyAlaValAlaGlyTyrrIlePheGlyLeuPheLeuArgIleThrGlyGly 450
DB 1315 AACGTCGCGGTGCTGTAGATGGAATGTTTCAGGTGCTTAATCGTTTGG--- 1371
QY 451 GluProTyrrLeuTyrrLeuGlnPheProLeuIlePheTyrrProGlyTyrrTyrrSerAspLysAsn 470
DB 1372 -----ATTGCATGATTAACCATTTG-----GCACATATCAACGAAATATTC 1413
QY 471 GlyIleTyrrAsnGlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThr 490
DB 1414 GCGTTATAT-----GAAATTAATCTCGGATTTATTGTA 1446
QY 491 AsnIleCysValSerTyrrLeuAlaLysTyrrLeu 501
DB 1447 AGGTATATCGTTACATATGTGTGAAGTAAACTT 1479

RESULT 6
US-11-266-748A-56747
; Sequence 56747, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
```

```

; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56747
; LENGTH: 3227
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-56747

Alignment Scores:
Pred. No.: 1,84e-16 Length: 3227
Score: 245.50 Matches: 125
Percent Similarity: 36.2% Conservative: 79
Best Local Similarity: 22.2% Mismatches: 162
Query Match: 8.2% Indels: 197
Gaps: 27
DB:

US-10-724-806-4 (1-580) x US-11-266-748A-56747 (1-3227)

QY 4 HisValGlnGlyLeuValAlaIleIleLeuPheTyrrLeuLeuIlePheLeuValGlyIle 23
DB 105 CACGCTACGACATACGACGCTGTGTGTCATCTTGTCTGCTCATTCCTGTGGGATC 164
QY 24 TrpAlaAlaIleTyrrPheThrLysAsnSerGlyAsnProGluGlnArgSerGlnAlaIleIle 43
DB 165 TGGTGTCTCATCTGTCGA-----AGTGCAGGACCAATT 197
QY 44 ValGly-----GlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThr 59
DB 198 GCGCGTATTCTCTGCGCGGAGGTCACAGCTGTGGCCATTCGACATCTCGATG 257
QY 60 AlaThrTrpValGlyGlyGlyTyrrIleAsnGlyThrAlaGlnAlaValTyrrGlyProGly 79
DB 258 TCACGCAATGTGGGACGAGTGTGTTTCATCGGCTGGCT-----GGGACAGGG 305
QY 80 Cys-----GlyLeuAlaIleTrp----- 84
DB 306 GCTGCCGAGGCTTCGCTGAGTGTGCTTCGATGACATAGCAATCAAGCTCGA 365
QY 85 -----AlaHisAlaProIleGlyTyrrSer----- 92
DB 366 GGAAGACAGAGGATCCATCCAGATCCAGTCACGAGAGACTGGGTCGTCGCCAGCAGACC 425
QY 93 ---LeuSerLeuIleLeuGlnGlyLeuPhePheAlaLysProMetArgSerLysGlyTyrr 111
DB 426 TGGCTGCTCTCGCCCTTGGCTGTGCTTCGCTTCCTGTCATATC---GCAGAGGTGTG 482
QY 112 ValThrMetLeuAspProPheLysGlnIleTyrrGlyLysArgMetGlyGly----- 128
DB 483 GTCACAATG-----CCGCAATCTCGAAGACGATTTGGGGGCCAGAGAGAT 530
QY 129 -----LeuLeuPheIleProAlaLeuMetGlyGluMet 139
DB 531 CAGGTATACATGCTGTCTGTCTCTCATCTTCATCTTCACCAAGATCTCGACTGAC 590
QY 140 PheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspVal 159
DB 591 ATCTTCTCTGAGGCTCTTC-----ATCCAGATGCAATTTGGGCTGG 632
```

```
QY 160 AspyValaenileSerValIleValSerAlaLeuIleAlaIleLeuTyrThrIleuValGly 179
DB 633 AACCTGACCTCTCCACAGAGCTGCTGGTGGAGCTGCGCTTACACCATTCAGAGT 692
QY 180 GlyLeuTyrSerValAlaTyrThrAspValValGln----- 191
DB 693 GGCCTCATGGCGGTATCTACACAGATGCTCTGACAGCGTGATCATGTAGGGGAGCC 752
QY 192 LeuPheCysIlePheIleGlyLeu-----Tyr----- 200
DB 753 CTGGTCCCATGTTCTGGGCTTTAGAGACGTGGCTGTACCCAGGCTGAGAGACGG 812
QY 201 -----IleSerValProPheAlaLeuSerHis----- 209
DB 813 TACAGGAGGCCATCCCTAATGTCACAGTCCCAACACCACCTGTACCTCCAGCGCC 872
QY 210 -----ProAlaValThrAspIleGlyPheThrAlaValHis 221
DB 873 GATGCTTCCACATTCTTGGGACCTGTGAGCGGGGACATC----- 914
QY 222 AlaIlyTyrGlnSerProTyrPheGlyThrIleGlnSerValGluValTyr---ThrTyr 240
DB 915 -----CCTTGGCAGGTCTCATTTTGGGCTCAGAGTGTGGCCACCTGG 959
QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTyr-----GlnAlaTyr 257
DB 960 TGT-----TGGTGACACAGACCGTCAATT 983
QY 258 PheGlnArgValIleuSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAla 277
DB 984 GTGACGCGGTCTCTCGGCCAAGAGTCTGTCTCATGCAAGAGGAGGCTCGTGGTGGG 1043
QY 278 AlaPhe-----GlyCysLeuValMetAla 285
DB 1044 GGCTACCTGAAGATCTCCCATGTTCTTCATGTCATGCGCCGATATATCAGCGGGCC 1103
QY 286 Leu-----ProAlaIleCysIleGlyAlaIleGlyAla 296
DB 1104 CTGTCCACAGAGAGTGGCGTGCCTGAGACCTGATGTCCTGCAAGATCTGTGGGGC 1163
QY 297 SerThrAspTyrAsnGlnThrAlaTyrGlyTyrProAspProIlyThrIleGluAla 316
DB 1164 CGAGTGGGATGTTCCAACTTCCCTAC----- 1190
QY 317 AspMetIleLeuProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGly 336
DB 1191 -----CCTAAGTGGTCAAGGCCCTCATGCTGTGTGCTGGGGGGCTGATG 1238
QY 337 LeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSer 356
DB 1239 ATTGCCGTGATCATGCGCTCTCATGAGCTCCTCATCTCATCTTCAACGAGCGAGC 1298
QY 357 SerMetPheAlaArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspIleGlu 376
DB 1299 ACCCTGTTCCACATTGATGTGTGCGAG--CGCTTCCGACAGAACTCAACAGACAGAG 1355
QY 377 IleValTyrPValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAla 396
DB 1356 CTGATGCGGTGGGCGAGAGTGTGTG--GTGTTC----- 1388
QY 397 LeuLeuThrIlyThrValTyrGlyLeuTyr----- 406
DB 1389 CTGGTGTTCATCAGCATC-----CTGTGATCCCATCATCCAAAGCTCAACAGTGGG 1442
QY 407 -----TyrLeuSerSerAspLeuValTyrIleIle-----IlePheProGln 420
DB 1443 CAGCTCTTCGACTATCATCAGGCTGTACCAAGTATCTGAGCCCAACCATTCACCGCTCTC 1502
QY 421 LeuLeuCysValLeuPheIleIlysglyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440
DB 1503 TTCTGCTGGCGCATCTTCTGCAAGAGGAGTCACAGAGCCGGAGCTTTCTGGGGCTCGTG 1562
```

```
QY 441 PheGlyLeu 443
DB 1563 TTGGGCTGTG 1571

RESULT 7
US-11-021-837-39/C
; Sequence 39, Application US/11021837
; Publication No. US20060140972A1
; GENERAL INFORMATION:
; APPLICANT: Alm, Richard
; APPLICANT: Manning, Paul
; APPLICANT: McLaughlin, Robert
; APPLICANT: McCormack, Kathleen
; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
; FILE REFERENCE: 101324-US
; CURRENT APPLICATION NUMBER: US/11/021,837
; PRIOR FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/533534
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/600680
; PRIOR FILING DATE: 2003-08-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39
; LENGTH: 95963
; TYPE: DNA
; ORGANISM: Staphylococcus saprophyticus
US-11-021-837-39

Alignment Scores:
Pred. No.: 4,98e-13 Length: 95963
Score: 235.00 Matches: 124
Percent Similarity: 40.1% Conservative: 102
Best Local Similarity: 22.0% Mismatches: 249
Query Match: 7.9% Indels: 88
DB: 7 Gaps: 21

US-10-724-806-4 (1-580) x US-11-021-837-39 (1-95963)
QY 9 ValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyIleTyrAlaIleTyrIly 28
DB 55486 ATAAATGATCATCATATATATTTACATATATATATATATATATATATATATATATAT 55433
QY 29 ThrIlyAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyIlyArgAsp 48
DB 55432 AAACAGCTACAGGTAATTAAGTGA-----TTATAGTTAGGTGAGGAGC 55385
QY 49 IleGlyLeuLeuValGlyIlyPheThrMetThrAlaThrTyrPValGlyIlyTyrIle 68
DB 55384 ATCGGTCTTATGTGACAGCACTTTCACAGAGTGCATGATGAGTGGCTGATGATT 55325
QY 69 AsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTyrAlaHisAlaPro 88
DB 55324 ATGGGTCAACAGAGTCTGTTATAGTACTGGTATATCTGCTANGTG-----ATTACC 55271
QY 89 IleGlyTyrSerLeuSerLeuIleLeuGlyIlyLeuPhePheAlaIlyPrometArgSer 108
DB 55270 ATGTGCTTTCAATAGGCGCTATGTAAT-----TATTTGTGCTGACCTGTGCTA 55217
QY 109 LysGlyIly-----ValThrMetLeuAspProPheIleGlnIle 121
DB 55216 CGTGTCTATACAGAAATTTGGCTGGTATGCGGATTCATTCACAGATTTCTTTAAAAACA 55157
QY 122 TyrGlyIlyArgMetGlyIlyLeuLeuPheIleProAlaLeuMetGlyIlyMetPheTyr 141
DB 55156 TTAATATGACCAATATATATATATATATATATATATATATATATATATATATATATAT 55097
QY 142 AlaAlaAlaIlePheSerAlaLeu-----GlyAlaThrIleSerValIleIleAsp 158
DB 55096 ACTTTATACACACATTCGGGTTTGTATCAGGTGTAATATATATATATATATATATATAT 55037
QY 159 ValaPheValaenileSerValIleValSerAlaLeuIleAlaIleLeuTyrThrIleuVal 178
DB 159 ValaPheValaenileSerValIleValSerAlaLeuIleAlaIleLeuTyrThrIleuVal 178
```

Db 55036 TTGAATTATCATGGGATTAAGTGTGATTTATGTCATTTATTAACATTTTC 54977  
 QY 179 GYGLIleuTyrSerValAlaTyrThrAspValAlaGlnleuPheCysIlePheIleGly 198  
 Db 54976 GGTGGTATTCCTGCTATCTATCACCAGCTTCTTCCAAAGAGTATGATTTAAAGCGCT 54917  
 QY 199 LeuTyrIleSerValProPheAlaLeu-----SerHisPro 210  
 Db 54916 ATG---GTATGGTCCCTTATTCGTTCGATTCATTCATTAAGCGCATTCATGCTTTAA 54860  
 QY 211 AlaValThrAspIleGlyPheThrAlaValHisAlaIleGlyGlnSerProTyrIleGly 230  
 Db 54859 CAAGTGCAGAGATGAACCAACCAATATGATCTTTTAAAGCACAACCGTTTAAAGT 54800  
 QY 231 ThrIleGlnSerValGluValTyrThrTyrPheLeuAspAsnPheLeuLeuMetLeuGly 250  
 Db 54799 ATCAAT-----TCAGTGTTCCTGG-----GGATTGGGT 54770  
 QY 251 GlyIleProTyrGlnAlaTyrPheGlnArgValLeuSerSerSerAlaThrTyrAla 270  
 Db 54769 TATTTTCGCAACCTCATATTTATGTTAGTTATGCTATAAATCGCAACAATTATTA 54710  
 QY 271 GlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeuProAlaIleCys 290  
 Db 54709 CCAAAAGCAGCAAGATTAGTATTAGCTGATGATGATGCTGTTATAGGCTGTTGCA 54650  
 QY 291 IleGlyAlaIleGlyAlaSerThrAspTyrAsnGlnThrAlaTyrGlyTyrProAspPro 310  
 Db 54649 GTTGGTTTAAACAGATTCGCTTT---ATTTCTGAACGCAACATCAAGTTAGAGATCC 54593  
 QY 311 LysThrIleGlnGluIleAspMetIleLeuProIleValLeuGlnTyrLeuCysProVal 330  
 Db 54592 GAAACT-----TTATTTATGCTTATGAGCCAAATTTATTCACCCCT 54551  
 QY 331 TyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAspSer 350  
 Db 54550 TTAGTTGGAGGCTTCTTATTTGGCAGCAATTTTGGCAGCCATCATGATGATCAATTTCTTCA 54491  
 QY 351 SerIleLeuSerIleSerMetPheAlaArgAsnIleTyrIleLeu-----Ser 367  
 Db 54490 CAATTACTGCTTCACTTCAAGTCTCTTAACGCAAGACTTTTAAACCTTCCTGCTGTA 54431  
 QY 368 PheArgGlnAsnAlaSerAspLysGluIleValTyrValMetArgIleThrValLeuVal 387  
 Db 54430 GATTAAGCTAAAGCACAACGAGAAAGATCTTAATGTTGCTGCTTATCTGATTAAT 54371  
 QY 388 PheGlyAlaSerAlaThrAlaMetAlaLeuThrIleValTyrGlyLeuTyr 407  
 Db 54370 GTAGCGATTGCGCTATGATGATGCT-----TGCTCG 54338  
 QY 408 LeuSerSerAspLeuValTyrIleIle-----Ile 417  
 Db 54337 CCAAAAGATACGATTAATTAATTTAGTTGATCTGGGAGAGATTTGGTCCGCAATTT 54278  
 QY 418 PheProGlnLeuLeuCysValLeuPheIleIleGlyThrAsnThrTyrGlyAlaValAla 437  
 Db 54277 AGTCCATTAAGTATCTTCTCATCTCTATGGAAGGCTATACGTAAGTGGTGA 54218  
 QY 438 GlyTyrIlePheGlyLeuPheLeuArgIleThrGlyGlyIleProTyrLeuTyrLeuGln 457  
 Db 54217 GGATGATTAACCGCCCATTTGCTGATTAATTTG-----ATTGTTGATTA 54167  
 QY 458 ProLeuIlePheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPhe 477  
 Db 54166 CCATTA-----GCTTATCATAGAGCTATTTGGTATGAT---GAAATTAAT 54122  
 QY 478 ProPheIleThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTyrLeu 497  
 Db 54121 CCTGGAATTCCTTACAGATGATTAATTAATTT-----GTT 54083  
 QY 498 AlaIleTyrIleuPheGlnSerGlyThr---LeuProTyrIleuAspValPheAspAla 516  
 Db 54082 AGTAATAATCACTAAACCTGGTTCATTTGTGACACATGACCTAGATAAAGTAAACAA 54023

QY 517 ValValAlaArgHisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsn 536  
 Db 54022 ATCGTA-----AAGGAATGTTGATATTAATTAAGCTGTTAAATCAATACAAAC 53975  
 QY 537 IleIleLeu 539  
 Db 53974 GCCAAACTG 53966  
 RESULT 8  
 US-11-217-529-1082  
 / Sequence 1082, Application US/11217529  
 / Publication No. US2006009612A1  
 / GENERAL INFORMATION:  
 / APPLICANT: SUNTORY LIMITED  
 / APPLICANT: NAKAO, YOSHIIRO  
 / APPLICANT: NAKAMURA, NORIHISA  
 / APPLICANT: KODAMA, YUKIKO  
 / APPLICANT: FUJIMURA, TOMOKO  
 / APPLICANT: ASHIKARI, TOSHIIKO  
 / TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
 / FILE REFERENCE: S-38-285  
 / CURRENT APPLICATION NUMBER: US/11/217,529  
 / CURRENT FILING DATE: 2005-09-02  
 / PRIOR APPLICATION NUMBER: US 10/932,182  
 / PRIOR FILING DATE: 2004-09-02  
 / NUMBER OF SEQ ID NOS: 197023  
 / SOFTWARE: PatentIn version 3.3  
 / SEQ ID NO 1082  
 / LENGTH: 2211  
 / TYPE: DNA  
 / ORGANISM: *Saccharomyces pastorianus*  
 US-11-217-529-1082  
 Alignment Scores:  
 Pred. No.: 1,95e-12 Length: 2211  
 Score: 206.50 Matches: 126  
 Percent Similarity: 36.3% Conservative: 87  
 Best Local Similarity: 21.5% Mismatches: 245  
 Query Match: 6.9% Indels: 129  
 Gaps: 23  
 US-10-724-806-4 (1-580) x US-11-217-529-1082 (1-2211)  
 QY 39 SerGluAlaIleIleValGlyArgAspIle-----GlyLeuLeuValGlyGlyPhe 56  
 Db 133 GCTGAAGAAATTTACCACTGCCGCTAGATCGGTAAAGACCGGTTAGTGGCGGCTGTG 192  
 QY 57 ThrMetThrAlaThrTyrValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyr 76  
 Db 193 GTTTCAGTTGGATTTGGTGTCTAGCTGTGTTGACTTATCCCAAGAAATACGAGAT 252  
 QY 77 GlyProGlyCysGlyLeuAlaTyrPalahHisAlaProIleGlyTyrSerLeuSerLeuIle 96  
 Db 253 GGATATTTGGTGGGATACCATATGCC-----GCCGGTCTGTTTCCAAATTAAT 303  
 QY 97 LeuGlyGlyLeuPhePheAlaLysProMetArg---SerLysGlyTyrValThrMetLeu 115  
 Db 304 GCGTTGCCATCTTGCTGCTTAAAGACCAAGCAATGCTCCCATCGCACACATCTCTG 363  
 QY 116 AspProPheIleGlnIleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeu 135  
 Db 364 GAATTAAGTGAAGAAAAGATATGTTAGTGGCCATGATGCTGATCTGTTTACGCCATT 423  
 QY 136 MetGlyGluMetPheThrPalahAlaIlePheSerAlaLeuGlyAlaThrIleSerVal 155  
 Db 424 GCAACCAATTTTGGTCACTTCTATGCTTTTAACTTCAAGGTCGCTGTTTCAAGTAC 483  
 QY 156 IleIleAspValAspValAsnIleSerValIleValSerAlaLeuAlaIleLeuTyr 175  
 Db 484 TTGACCGGATGAACCACTTGATCTTCTTCTACTACCGCTGCTGTGCTTAT 543  
 QY 176 ThrLeuValGlyGlyLeuTyrSerValAlaTyrThrAspValAlaGlnleuPheCysIle 195



```

QY 136 MetGlyGluMetPheThrAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerVal 155
   |||
Db 424 GGGAGGATATTTTGTGCTCATGCTTTAACTTCAGGTTCTGCTTCTTAACTGAT 483
QY 156 IleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuYr 175
   |||
Db 484 TTAACCGGATGAACATGCTATGCTGTTTAACTGCTGCTGGGCTGTTGTTTAT 543
QY 176 ThrIleuValGlyGlyLeuYrSerValAlaYrThrAspValValGluIleuPheCysIle 195
   |||
Db 544 ACTCAATTTGGTGTATTAAGCACTTTCTTAAGCGACTATATGCAACATGTCAT 603
QY 196 PheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIle 215
   |||
Db 604 ATCATC-----ATTGCTCTCGATTTGCTTTAAAGTTTAAGCTTACGATGATGTT 654
QY 216 GlyPheThr-----ATTTGCTCTCGATTTGCTTTAAAGTTTAAGCTTACGATGATGTT 654
   |||
Db 655 TTAGGCTCACCGGAAAAGTTTATGACTTATGCTTATGCTGAACCGCCAAAGGATCCAGTA 714
QY 221 HisAlaIleYrThrGlnSerProTrpLeuGlyThrIleGluSer-----Val 235
   |||
Db 715 GACGGTAACTATCAAGTGAATATATG--ACCATGACATCCAAATCCCGTGTATTTTA 771
QY 236 GluValIleYrThrTrpLeuAspAsnPheLeuLeuMetLeuGlyGlyIleProTrpGln 255
   |||
Db 772 TTAATTTATTAACCTGATTTGGAATTTTCCGACCGTTTCTTGAT-----AAT 819
QY 256 AlaIleYrPheGlnArgValLeuSerSerSerSerAlaThr-----TyrAlaGlnValLeu 273
   |||
Db 820 GGTATTTGGAATTAAGCGATTTCTGCTAGTCCCGACGAGTTTGAACATATATGCGATC 879
QY 274 SerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAla 293
   |||
Db 880 GGTGGGTATGACATGTTTCA-----GTAACCTTTTATTTATTTCAATGACATGATTA 933
QY 294 IleGlyAlaSerThrAspTrpAsnGlnThrAlaIleYrGlyIleProAspProIleYrIle 313
   |||
Db 934 GCATGCTTGGCGGTGAACAGTCTCCAACTTCCCACTATCTGATCCACTTACTTGG 993
QY 314 GluGluAlaAsp-----MetIleLeuProIleValIleuGlnIleuYrCysProValIle 331
   |||
Db 994 TTCCAGGCAAAATTTCTGCTGTTGCTTCCGCGACGCTGA-----1032
QY 332 IleSerPhePheGlyLeuGlyAlaValSerAla-----AlaVal 344
   |||
Db 1033 ATTGCTATCATGGGTAAAGGGGGTCTGTCGCACTGCTATATGATTTTCAATGCGCTC 1092
QY 345 MetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleYr 364
   |||
Db 1093 ACATCTGCTATGCTGCTGCACTGATGCGCTTCACTGCTTTCATCACTTAACATATCTAT 1152
QY 365 GlnLeuSerPheArgGlnAsnAlaSerAspIleGlyIleValIleYrValMetArgIleThr 384
   |||
Db 1153 AGAGAAATATATGATCTCTGTCGCAAGCGGTGAAGAAATGATTTAATCAATCACACGTTGCT 1212
QY 385 ValIleValPheGlyAlaSerAlaThrAlaMetAla-LeuLeuThrIleYrThrValIle 404
   |||
Db 1213 TGTATCTTTTGTGCTTCCATGAGTGAATTTTCCGTTGGTTTAACTATGATGCTGATTT 1272
QY 404 YleuTrpIleLeuSerSerAspLeuValIleIleIlePheProGlnLeuLeuYrCysVal 424
   |||
Db 1273 TCTATGCTTATATCTAT-GAATGATGGGTAAATATTATTAAGTACGCTATATCTGCT 1331
QY 424 LeuPheIleYrGlyThrAsnThrIleYrGlyAlaValAlaGlyIleYrIlePhe 441
   |||
Db 1332 CGTTTGGACCTTATGTTCCAAACATGATTTGGTGGCGCTGTAGTGTGCTATTTT 1391
QY 441 eGlyLeuPheLeuArgIleThrGly-----449
   |||
Db 1392 GGGCACAGACGCTGCTATATGCTATGCGCTTGTCTGTACCAATCCCTTTATATAAGATTT 1451
   |||
QY 450 -----GlyGluProTrpIleuYrIle 456

```

```

Db 1452 GACCGTGATACCTACGTTATCATGATTTATCAATGTTAAACAGGTACTTGGCTTGTGCT 1511
QY 456 uGlnProIleuIlePheYrProGlyIleYr-----TyrSerAspIleAsnGlyIleYrAsnG 475
   |||
Db 1512 ATCACCAGCATTTTATTTATCTATTTTAAAGTATGCTTTTAAAG-----CCACA 1559
QY 475 nArgPheProPheIleYrThrLeuSerMetValThrSerPhePheThrAsnIleCysValSe 495
   |||
Db 1560 AAATTTTGACTGGGAGAAATGAAGATATTAAGTAA-----GTGCA 1601
QY 495 rTyrLeuAlaYrIleuPheGluSerIleYrLeuProPolysIleuAspValPheAs 515
   |||
Db 1602 CGAACTCGACAGATGATTT-----CAGCGTCACTCGTATATCCAGCTTACGA 1649
QY 515 PalValValAlaAlaAsnHisSerGluAsnMetAspIleThrIleLeuVal-----532
   |||
Db 1650 TGCT--GAAAGTAAAGTATGAGTAAAGGAGAGAGAAACAAATTTCTGTGCTTCAGATAG 1706
QY 533 -ArgAsnGluAsnIleYrLeuAsnGluLeuAlaProValIleProArgGlnSerLeuThr 552
   |||
Db 1707 TGAATAAAACGATGTTAGAGTAAATATGAAATTAATGATTTGAGCTTAACCTGCTGTGCT 1766
QY 552 rLeuSerSerThrPheThrAsnIleGlu 561
   |||
Db 1767 GATTAAGTAAAGCATTGTTTCAAGAAAGAT 1794

```

RESULT 10

US-10-471-571A-1991

Sequence 1991, Application US/10471571A

Publication No. US20060115490A1

GENERAL INFORMATION:

APPLICANT: CHIRON SPA

TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE: P026927WO

CURRENT APPLICATION NUMBER: US/10/471,571A

CURRENT FILING DATE: 2003-09-12

PRIOR APPLICATION NUMBER: GB-0107661.1

PRIOR FILING DATE: 2001-03-27

NUMBER OF SEQ ID NOS: 5642

SOFTWARE: SeqMan99, version 1.03

SEQ ID NO 1991

LENGTH: 1539

TYPE: DNA

ORGANISM: Staphylococcus aureus

US-10-471-571A-1991

Alignment Scores:

Pred. No.:	1,08e-11	1539
Score:	197.50	118
Percent Similarity:	37.3%	Conservative: 104
Best Local Similarity:	19.8%	Mismatches: 234
Query Match:	6.6%	Indels: 129
DB:	6	Gaps: 23

US-10-724-806-4 (1-580) x US-10-471-571A-1991 (1-1539)

```

QY 9 ValAlaIleIleLeuPheYrIleuIlePheLeuValGlyIleThrAlaIleTrpIle 28
   |||
Db 43 GTTCCGCTTATATATCTATCTAGCTATGCTGTTGCTTATGCTTATTTTAAACAGCG 102
QY 29 ThrIleAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyIleArgAsp 48
   |||
Db 103 GCGAGCCAAAGTAAACAT-----AGTTCTTAAACCGCAAGTGTGCTC---144
QY 49 IleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyIleYrIle 68
   |||
Db 145 TTGCCATCTTGGGTAGTGGCTTTTCAATTTATGCTACACCTTAAGTACGATTAACATTT 204
QY 69 AsnGlyThrAlaGluAlaValIleYrGlyProGlyCysGlyLeuAlaThrAlaHisAlaPro 88
   |||
Db 205 ATGTGACACCGAGAAAGCATTTTAAACAGAT-----TGTCTATATATGCT 252

```

QY	89	IIlegIYrSerSerLeuLeuLeuGlyGlyLeu---PhePhAlaIysProMetArg	107
Db	253	GGTAACATTGCATATCGCCCAATTATTCATTACTTATTTATTTCTATGTCCTTTCTTT	312
QY	108	SerLYsgLYrValThrMetLeuAspProPheLYglnIle--TYrGLY-----	123
Db	313	AAAAAGTTAAAGTTAACTGCATCTGCATATGCAATTTTAAAGACATAGATTTGGCCCTTCACATA	372
QY	124	LysArgMetGLYLeuLeuPheIleProAlaLeuMetGLYGlutMetPheTrpAlaIa	143
Db	373	CGTGCATTGGCTCATTTATTTATTTGCTTTACCATTTTGGCGCGTGT-----GCAATT	426
QY	144	AlaIlePheSerAlaLeuGlyValThrIleSerValIleIleAspValAspValAsnIle	163
Db	427	GTTATCTACTTA-----CCAACTATGACATACATCTGTATTCAGACATGAACCT	477
QY	164	SerValIleValSer-----AlaLeuIleAlaIleLeuTYrThrLeuValGlyGlyLeu	181
Db	478	TATATCGGTGCATCACTCGTGGTGTACTATATATTTTATTTATTAATCAATTTTATGATGGTTC	537
QY	182	TYrSerValAlaTYrThraspValValGlnLeuPheCysIlePheIleGlyLeuTrpIle	201
Db	538	GAAGGTGTGGTTTGGAGTATTCATTCAGAGCGCATTTATTTAGCGGCGCTTACTGT	597
QY	202	SerValProPheAlaLeuSerHisProAlaValThraspIleGlyPheThralaValHis	221
Db	598	-----ATTATATTTCTAGGCGTGTGACAACTTAAA	627
QY	222	AlaLYrTYrGlnSerProTrpLeuGlyThrIleGlnSerValGlnValTYrThrTrpLeu	241
Db	628	GGCGGTTTGGCAGCTGCTTTGCAATGCGATTTGACACAAAAATTAATTAGT--GCA	684
QY	242	AspAsnPhelLeuLeu-----LeuMetLeuGlyGlyIle	252
Db	685	GACAAATTGGAACCTAAATACTCGCGACATGCCATTCCAAATATTTTCTAGAAATATT	744
QY	253	-----ProTrpGlnAlaTYrPheGlnArgValLeuSerSer	264
Db	745	TTCAACAACTGTATCATATCACAGCGAGTCAAGAGCTGTCAGCGTATATCAACCTTCT	804
QY	265	SerSerAlaThrTYrAlaGln-----ValLeuSerPheLeuAla	277
Db	805	GATAGTTTAAAGAAACAATATAATGTTATGAGACAAATGATCTAGCTTATTAATTCA	864
QY	278	Ala-----PheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIle	294
Db	865	GCACCTTATTTATTTGGTATGGGTACAAAGCGGTATTCATTTTATACACATGACATCGTT	924
QY	295	-----GlyAlaSerThraspTrpAsnGlnThrAlaTYrGlyTYrProAspProLys	311
Db	925	TTACCAAAAGGCTTCAATCA-----	945
QY	312	ThrLYsgLInAlaAspMetIleLeuProIleValLeuGlnTYrLeuCysProValTYr	331
Db	946	-----TCATCTGTAGGCAATTTTCAATTTTACTAGACAGATGCCACATTT	990
QY	332	IleSerPhePheGlyLeuGlyAlaValSerAlaIaValMetSerSerAlaAspSerSer	351
Db	991	GTAGAGGAGTACTTATTCACAGCATTTTCCGCGGTGACAGCTTACCATTCATCTAGT	1056
QY	352	IleLeuSerAlaSerSerMetPheAlaArgAsnIleTYrGlnLeuSerPheArgIleAsn	371
Db	1051	TTAAATTTTATATTCGCTGTATTTTCATATTCACATTCAGCATTAAGCAACGCTTTCGGAAGAGT	1116
QY	372	AlaSerAspLYsgLInIleValTrpValMetArgIleThrValLeuVal-----	387
Db	1111	ACCGAGCGCACAGAGTTAACTTGCACGCTTCATTTATTTATTCATTCGACAGTATTTTCGGT	1176
QY	388	PheGlyAlaSerAlaThralMetAlaLeuLeuThrLYrThValTYrGlyLeuTrpLYr	407
Db	1171	TTTGAATCTACTATACTTATATGCTTCTAATTCAAATGACTTATGGAATTTATTTCTTG	1233
QY	408	LeuSerSerAspLeuValTYrIleIleIlePheProGlnLeuLeuCysValLeuPheIle	427

```

Db      1231 TTGTGACTGATTAATTCGGCGTTCCATTCGGCTGCTGATTAATTCGAGTTGGATTTTACT      ::::
Oy      428 LysGIyThrsanthrYrGIyAlaValAlAGLyTyrllePheglYleuPheleuAryIle      ::::
Db      1291 AAAGGTACGAATACATTCGGGTGATTAATTCGGATTAATATTTGGATATCATCTTT-----      ::::
Oy      448 ThrGIyGIyIuPProTyrlleuTYrLeuGIuPProleuIlePheTYrProGIyTYrTYrSer      ::::
Db      1345 -----GCTTATGCTAT-----      ::::
Oy      468 AspylsanGIyIleTYraenGIuAryPheProPhe-----LysThrlleuSerMetVal 485      ::::
Db      1357 -----AATGGTGTGGCAAAAGGTAATCATCTTTCTATGATTAATCATCAATTTCATTACA 1410      ::::
Oy      486 ThrserPhePheThrsanIleCysValSerTYrleuAlaLysTYrleuPheGIuSerGIy 505      ::::
Db      1411 GTTCCTTTTGCTTTGCTTTATATCATCTTATGCTCATTT-----      ::::
Oy      506 ThrlleuProPrylsleuAspValPheAspAlaValAlaArgHisSerGIuGIuAsn 525      ::::
Db      1447 -----GTCCTTCAAAACATATAAAAGATATA 1473      ::::
Oy      526 MetksplysThrlleuValArgsnGIuAsnIleLysleuAsnGIuLeuAlaProVal 545      ::::
Db      1474 ACGGGATTACCAATTTT-----GAA 1494      ::::
Oy      546 LysProArgIuSerIleuThrlleuSerSerThrlleuThrsanLys 560      ::::
Db      1495 AAAGATAAACCATCAACATACATTTCAAAAACGGCTACGAAAAAG 1539      ::::

RESULT 11
US-11-266-748A-26072
; Sequence 26072, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Hartkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcription Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662, 276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700, 293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26072
; LENGTH: 2984
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-26072

Alignment Scores:
Pred. No.: 1,566-10      Length: 2984
Score: 191.00      Matches: 120
Percent Similarity: 38.1%      Conservative: 70
Best Local Similarity: 24.0%      Mismatches: 197

```









QY 434 GlyAlaValAlaGlyTyrIlePheGlyLeu 443  
Db 2064 GGAGCTTCTGGGGCTCGTGTGTGGCCTG 2093  
RESULT 14  
US-11-266-748A-187885  
; Sequence 187885, Application US/11266748A  
; Publication No. US2006013463A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Mulligan, Patrick  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483396  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO: 187885  
; LENGTH: 1578  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-187885  
Alignment Scores:  
Pred. No.: 1.95e-08 Length: 1578  
Score: 168.00 Matches: 97  
Percent Similarity: 38.5% Conservative: 56  
Best Local Similarity: 24.4% Mismatches: 160  
Query Match: 5.6% Indels: 84  
DB: Gaps: 18  
US-10-724-806-4 (1-580) x US-11-266-748A-187885 (1-1578)  
QY 3 PheHisValGluGlyLeuValAlaIleLeuPheTyrLeuIlePheLeuValGly 22  
Db 300 TTCTTCATCATGAGACTATGTGTGTCCTCGCTGCTGCTTCTCTCTGTCATTTGG 359  
QY 23 IleTrrAlaIle-----TrrLysThrLysAsnSerGlyAsnProGluGluArgSer 39  
Db 360 CTCTACCATGCTTCTGTCGTGGCTGGCGGCATCTACTGTTGTGTGAG----- 404  
QY 40 GluAlaIleIleValGlyGlyArgAspIleGlyLeuLeuValGlyGlyPheThrMetThr 59  
Db 405 -----CTGCTGATGGCGGACCGCAAAATGGCTGCTCTCCGCTGGGCACTGCTGCTG 458  
QY 60 AlaThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGly----- 77  
Db 459 GCCACCTTCCAGTCAGCCGTGGCCATCTCGGTGCGGTGAGAGATATCCGATTTGGG 518  
QY 78 -----ProGlyCys-----GlyLeu-----AlaTrrAla 85  
Db 519 ACCCAATATGTTGCTCTGCTGGGCTGCTACTTCTGCGGGCTGCTGATACCTGCACATC 578  
QY 86 HisIaIrrIleGlyTyrSerLeuSerLeuIleLeuGly-----GlyLeuPhe 101

Db 579 TTCAATCCCGTTTCTTACCGCCTGATCTCACAGTGGCTATGATGACTGAGCTTGC 638  
QY 102 PheAlaLysProMetArgSerLysGlyTyrValThrMetLeuAspProPheLysGluIle 121  
Db 639 TTCAATATAACTGTGCGAGTGTGTGGAACTGTGACCTTCAATC-----TTTCAGATGGTG 692  
QY 122 TyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPheTrr 141  
Db 693 ATC-----TACATGGGAGTGTGCTTATGCTCTCCGTCATG----- 728  
QY 142 AlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspVal 161  
Db 729 -----CCTCAATGAGTACTGCTTGTGATCTG 758  
QY 162 AsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGlyLeu 181  
Db 759 TGGCTGTCCGTGCTGCGCCCTGGCATTTGTCGTACCGCTATACAGCTCTGGGTGGCTG 818  
QY 182 TyrSerValAlaTyrThrAspValGluLeuPheCysIlePheIleGlyLeuTrrPile 201  
Db 819 AAGCCCTCATCTGGACAGATGTTCAGACACTGTCATGTTCTCGGG----- 869  
QY 202 SerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal--- 220  
Db 870 -----CAGCTGGCAGTATTATCATCTGTGGGGTTCAGCCAGGTGGGC 908  
QY 221 -----HisAlaLysTyrGlnSerProTrrPhe 229  
Db 909 GGGTTGGGGCGTGTGGCCGTGGCTTCCAGACAGCCGCGCATCTCGGGTTGAGCTG 968  
QY 230 GlyThrIleGluSerValGluValTyrThrTrrPhe-----AspAsnPheLeu 245  
Db 969 GATCCAGACCCCTTGTGCGGACACCTTCTGGAACCTTGGCGGGTGTCTTTCATG 1028  
QY 246 LeuLeuMetLeuGlyGlyIleProTrrGlnAlaTyrPheGlnArgValLeuSerSerSer 265  
Db 1029 ATGCTCTCTTATACGGGGGTAAAC---CAGGCTCAGGTGACGGGTACTCATGTTCCCGC 1085  
QY 266 SerAlaThrTrrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAla 285  
Db 1086 ACGGAGAG---GCTGCTGCTCTCTCCGTATATGACAGTTCCTCCCTTCAGAGGTGTC 1142  
QY 286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrrPasnGlnThrAlaTyr 305  
Db 1143 CTC-----TGGTGGGCTGCTCTCATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184  
QY 306 GlyTyrProAspProLysThrLysGluGluAlaAspMetIleLeu---ProIleValLeu 324  
Db 1185 TACCAAGAGTATCCCATGAGCATTCACAGGCTCAGGCAAGCCCAAGCAGTTCCTCTG 1244  
QY 325 GlnTrrLeuCysProValTrrLysSerPhePheGlyLeu---GlyAlaValSerAlaAla 343  
Db 1245 TACTTTGATGATGATCTCTCAAGGCTTCAGGCTGCGCAGGCTCTTCAATGCTCGC 1304  
QY 344 ValMetSerSerAlaAspSerSerIleLeuSerAlaSerMetPheAla 360  
Db 1305 CTCTTCAGCGGCTCTCTCAGCATATATCTGCTGTTTATTCATTCAGCA 1355  
RESULT 15  
US-11-266-748A-16097  
; Sequence 16097, Application US/11266748A  
; Publication No. US2006013463A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Mulligan, Patrick  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2

```

? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105482.6
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105483.4
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105507.0
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105485.9
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: EP 04105484.2
? PRIOR FILING DATE: 2004-11-03
? PRIOR APPLICATION NUMBER: US 60/662,276
? PRIOR FILING DATE: 2005-03-14
? PRIOR APPLICATION NUMBER: US 60/700,293
? PRIOR FILING DATE: 2005-07-18
? NUMBER OF SEQ ID NOS: 483996
? SOFTWARE: PatentIn version 3.3
? SEQ ID NO 16097
? LENGTH: 936
? TYPE: DNA
? ORGANISM: Homo Sapiens
US-11-266-748A-16097

```

Alignment Scores:	
Pred. No.:	4,24e-06
Score:	13.50
Percent Similarity:	36.48%
Best Local Similarity:	22.5%
Query Match:	4.6%
DB:	8
Length:	933
Matches:	58
Conservative:	37
Mismatches:	78
Indels:	85
Gaps:	9

US-10-724-806-4 (1-580) X US-11-266-748A-16097 (1-936)

QY	5	ValGIGlyLeuValAlaIleIleLeuPheTyrlLeuIlePheLeuValGlyIleTrp	24
Db	70	GTGGCGACATCATCGCTCATCTACTGTGATTTTGTCTGAATGTGGCCGTGGCATATGG	1238
QY	25	AlAlaIleTrpLys---ThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIle	43
Db	130	TCCTCTTGTCGGGCGAGTGAACACACGGTGAAAT-----GGCTACTTC	1718
QY	44	ValGlyGlyArgAspIle-----	49
Db	172	CTGGACGGCGCGGACATGACGTGTGGCCGATTTGAGCGCTCTCTTCGCCAGCAGCGAG	2318
QY	50	-----GlyLeuLeuValGly	54
Db	232	GGCTCTGGCGCTCTTCAATTGAACTGGCGGGCTCAGCGCGGACGAGAGGTCTGGCGCTGCA	2918
QY	55	GlyPheThrMetThrAlaThrTrpValGlyGlyGlyTyrIleAsnGlyThrAlaGluAla	74
Db	292	GGCTTGGAGTGAATGCCACGTACGTG-----	318
QY	75	ValTyrGlyProGlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTyrSerLeuSer	94
Db	319	-----CTGCTGGCAGCTGGGATGGGTGTCTGTGCCCATCTACATCTCTCTCAGAG	366
QY	95	LeuIleLeuGlyGlyLeuPhePheAlaLys-----PrometArgSerGlyGlyTyr	1118
Db	367	ATGTGACCTTACTGATGACATTCAGAAAGCGCTACGGGGGCGACGGATCGCATGTAC	426
QY	112	ValThrMetLeuAspProPheLysGlnIleTyrGlyLys-----ArgMetGly	1278
Db	427	CTGTCTGTCTGTCCCTGCTACTGTCTGTCTTCCACCAAGATATGCTGGACCTGTACGG	486
QY	128	GlyLeuLeuPheIleProAlaLeuMetGlyGluMetPheTrpAlaAlaIlePheSer	1478
Db	487	GGGGCTCTGTTGTGTGACATCTGGCGGTGGCTGAACCTTCAC-----	5328
QY	148	AlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIleSerValIleVal	1678
Db	529	-----CTTCCACATCTC	5438

```

Oy 168 SerAlaLeuIleAlaIleLeuTyrThrIleValGlyGlyLeuTyrSerValAlaTyrTr 187
    ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 544 AGCTGGGCATACAGCCCTGTACACCCATCGCAGGTGGCCTGGCTGTATCTACAG 603
Oy 188 AspValValGlnLeuPheCysIlePheIleGlyLeuTyrPheSerValProPheAlaLeu 207
    ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 GACGCCCTGCAGACGCTCATCATGTTGGTGGGGGCTGTATC-----CTG 648
Oy 208 SerHisProAlaValThrAspIle---GlyPheThrAlaValHisAlaLeuTyr 224
    ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 649 ACAATCAAAAGCTTTTGACCCAGATCGTGTTACGGGCGACGTGGAGGCACCTTAC 702

```

Search completed: July 10, 2006, 14:56:38  
Job time : 503 secs

**This Page Blank (uspto)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:34 ; Search time 8716 Seconds  
(without alignments)  
11182.583 Million cell updates/sec

Title: US-10-724-806-3  
Perfect score: 1743  
Sequence: 1 atgccttcacatgtaag.....ctgaataattacaaga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4826798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

EST:  
1: gb\_esc1:\*  
2: gb\_esc3:\*  
3: gb\_esc4:\*  
4: gb\_esc5:\*  
5: gb\_esc6:\*  
6: gb\_esc7:\*  
7: gb\_esc8:\*  
8: gb\_esc9:\*  
9: gb\_esc10:\*  
10: gb\_esc11:\*  
11: gb\_esc12:\*  
12: gb\_esc13:\*  
13: gb\_esc14:\*  
14: gb\_esc15:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1731.8	99.4	4097	6 AK053063	AK053063 Mus muscu
2	1731.8	99.4	4306	6 AK034415	AK034415 Mus muscu
3	1465	84.1	1743	14 AY413300	AY413300 Mus sapi
4	1162.6	66.7	1743	14 AY413298	AY413298 Homo sapi
5	1162.2	66.7	1743	14 AY413299	AY413299 Pan trogl
6	851.8	48.9	4122	6 AK164116	AK164116 Mus muscu
7	603.4	34.6	707	5 CD350164	CD350164 UI-M-FY0-
8	576.4	33.1	669	4 BY727598	BY727598 BY727598
9	576.4	33.1	882	5 CK704440	CK704440 958613 MA
10	472	27.1	828	14 CT403380	CT403380 Sus scrof
11	409.6	23.5	650	7 BB626260	BB626260 BB626260
12	406.8	23.3	672	14 AG157499	AG157499 Pan trogl
13	398.4	22.9	548	9 DA361315	DA361315 DA361315
14	391.8	22.5	576	9 DN991044	DN991044 TC112868
15	389.6	22.4	516	7 BE233479	BE233479 139685 MA
16	370	21.2	675	4 BY729567	BY729567 BY729567
17	367	20.1	988	10 DM665738	DM665738 CNB333-A0
18	359.6	20.6	781	10 DT197042	DT197042 JGI_CAA85
19	359.6	20.6	843	10 DT207246	DT207246 JGI_CAA81

20	356.4	20.4	1114	10 DT972151	DT972151 CLJ138-H0
21	348.8	20.0	753	10 DV838598	DV838598 LB01121.C
22	325.6	18.7	874	5 CK449221	CK449221 892772 MA
23	302.2	17.3	834	9 CX840109	CX840109 JGI_CAA9
24	301.6	17.3	541	7 AM668962	AM668962 111664 MA
25	301.6	17.3	1054	8 DM614674	DM614674 CLJ282-B0
26	273	15.7	694	8 CX033587	CX033587 1345948 N
27	263.8	15.1	524	7 BE723927	BE723927 198406 MA
28	252.6	14.5	800	1 AL669749	AL669749 AL669749
29	251.2	14.4	652	7 BB626456	BB626456 BB626456
30	236.2	13.6	658	8 CX196909	CX196909 SA_mxo_52
31	235.8	13.5	605	3 BQ829470	BQ829470 LG612176
32	229	13.1	549	3 DM274870	DM274870 BM274870
33	228.8	13.1	1021	10 DM616805	DM616805 CLJ294-F1
34	226	13.0	541	5 CJ145625	CJ145625 CJ145625
35	225.4	12.9	941	5 CD360297	CD360297 AGENCOURT
36	220.6	12.7	565	9 DA911919	DA911919 DA911919
37	215.8	12.4	658	3 BM629925	BM629925 17006875
38	215.4	12.4	641	2 BI630566	BI630566 RH59836.5
39	215	12.3	574	9 DA366634	DA366634 DA366634
40	213.6	12.3	640	2 BI629504	BI629504 RH58381.5
41	212.8	12.2	598	8 CX196517	CX196517 SA_mxo_46
42	212.4	12.2	1037	1 AL666817	AL666817 AL666817
43	209.8	12.0	682	9 DN604180	DN604180 ACAC-aab8
44	203.6	11.7	869	9 CX809396	CX809396 JGI_CAAJ1
45	200	11.5	583	3 BM277281	BM277281 BM277281

#### ALIGNMENTS

RESULT 1  
AK053063  
LOCUS  
DEFINITION  
AK053063 4097 bp mRNA linear HTC 02-SEP-2005  
Mus musculus 15 days embryo head cDNA, RIKEN full-length enriched library, clone:DJ03038E20 product:solute carrier family 5 (cholesterol transporter), member 7, full insert sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK053063  
AK053063.1 GI:26343192  
HTC, CAP trapper.  
Mus musculus (house mouse)  
Mus musculus

REFERENCE  
AUTHORS  
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. High-efficiency full-length cDNA cloning

JOURNAL  
PUBMED  
Meth. Enzymol. 303, 19-44 (1999)  
10349636

REFERENCE  
AUTHORS  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL  
PUBMED  
Genome Res. 10 (10), 1617-1630 (2000)  
11042159

REFERENCE  
AUTHORS  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tachito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL  
PUBMED  
Genome Res. 10 (11), 1757-1771 (2000)  
11076861

REFERENCE  
AUTHORS  
4 THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection

JOURNAL REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	The PANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the PANTOM Consortium.
TITLE	Antisense transcription in the Mammalian Transcriptome
JOURNAL REFERENCE	Science 309, 1564-1566 (2005)
AUTHORS	7 The PANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
TITLE	The Transcriptional Landscape of the Mammalian Genome
JOURNAL REFERENCE	Science 309, 1559-1563 (2005)
AUTHORS	8 (bases 1 to 4097) Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fuyuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Horii,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Komori,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sobabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
TITLE	Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsunumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers
COMMENT	1..4097 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:D930038E20" /db_xref="taxon:10090" /clone="D930038E20" /tissue.type="head" /clone_lib="Riken full-length enriched mouse cDNA library" /dev_stage="15 days embryo" 512..2254 /note="unnamed protein product; putative solute carrier family 5 (choline transporter), member 7 [MG]  GI:1927126 GB NM_022025, evidence: BLASTN, 99%, match=1743") /codon_start=1 /protein_id="BAC35253.1" /db_xref="gi:26345193" /translation="MSFVEVGLVAIIIFLLPIPLVGIAWAAKTNSGNPERSSRAIIV GGRDIGLVGFETATVWGGYNGITAEAYGCGCLAMQAADISLSILIGLFF AKPMRSQVTMLDPFOOI YGRMGLGFI PALNGEMMAAASAGATSVIIDVD VNISVSALAIILTYTVIGLVSYAVDVOLFCIFLGIMTSVPALSPADVIGFT AVAHKYSOPMGTISEVRYVYLINFLIMLGVIPKPMXYPORVYSSSATYAQTSELAAPFGCLWVALPAICGAISTDMNQRTAYGPDEPTKEABMWILLPVLYQYCPYISF FGDAVASAAVMSSADSSITLSAWSMFARNIYOLFSRQNASREIWMKLTIVLVGASGA TMMALLTKTVYGLWYLSDDIVYIIIFIPLLCLVFIRKNTYGAAGVYFGLEPITGSGA
FEATURES	source
CDS	

ORIGIN	EXPIXYLOPLFVGGYVSDKNGIINORPEPTLSMVNVSFPFNICVSYAKLIFPSGLT PKLDVDFAVARHSBEMDKTIIIVRENITKLNELAPKPROSLTSLSTPFNKREALDV DSSPSGSGEDNLO"
Query Match	99.4%; Score 1731.8; DB 6; Length 4097;
Best Local Similarity	99.6%; Pred. No. 0;
Matches 1736;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	1 ATGCGTTTCCATGTGGAAGACTGGTAGCTATTATCTCTTCACTCCTTATATTTCTG 60
DB	512 ATGCTTTCCACGTAGAAGACTGGTAGCTATTATCTCTTCACTCCTTATATTTCTG 571
QY	61 GTTGGAAATATGGCTGCATGGAACCAAAAAACGGGGCAACCCAGAAAGCCGAGTGA 120
DB	572 GTTGGAAATATGGCTGCATGGAACCAAAAAACGGGGCAACCCAGAAAGCCGAGTGA 631
QY	121 GCCATCATAGTCGGGGGCCCGTGACATTGGTTGGTTGGTTGGTTTACATGACAGCC 180
DB	632 GCCATCATAGTCGGGGGCCCGTGACATTGGTTGGTTGGTTGGTTTACATGACAGCC 691
QY	181 ACCTGGGTGGAGAGGCTCATCATATGGGACAGCAGAACGCTATATGGGCCAGTTGT 240
DB	692 ACCTGGGTGGAGAGGCTCATCATATGGGACAGCAGAACGCTATATGGGCCAGTTGT 751
QY	241 GGTCTAGCTTGGGCTCAGACCCCATGGATTCCTGTGAGTCTTAATTTTAGAGGCTG 300
DB	752 GGTCTAGCTTGGGCTCAGACCCCATGGATTCCTGTGAGTCTTAATTTTAGAGGCTG 811
QY	301 TTTTGTGGAAACCTATAGCGTTTCCAGAGGATATGTACTATGTATGACCATTAAACAG 360
DB	812 TTTTGTGGAAACCTATAGCGTTTCCAGAGGATATGTACTATGTATGACCATTAAACAG 871
QY	361 ATCTATGGAAGAGCCCATATGGGTGGGCTGCTTCTTCACTCCCTGCATGATGGGAGATGTC 420
DB	872 ATCTATGGAAGAGCCCATATGGGTGGGCTGCTTCTTCACTCCCTGCATGATGGGAGATGTC 931
QY	421 TGGGCTGAGCAATTTCTCTGCATTTAGGGGCCACCATACGCGATCATTTGATGTGGAT 480
DB	932 TGGGCTGAGCAATTTCTCTGCATTTAGGGGCCACCATACGCGATCATTTGATGTGGAT 991
QY	481 GTGGAACATATGCGATCATTTGCTCTGCACATCATTTGCCATTCTTTATACCTTATGAGTGGG 540
DB	992 GTGGAACATATGCGATCATTTGCTCTGCACATCATTTGCCATTCTTTATACCTTATGAGTGGG 1051
QY	541 CTCTACTCTGTGGCAATATCATGATGTTGTCCAGCTATTTCTGCATTTTATATGACCTGTG 600
DB	1052 CTCTACTCTGTGGCAATATCATGATGTTGTCCAGCTATTTCTGCATTTTATATGACCTGTG 1111
QY	601 ATCAGTGTCCCTTTTGGCCCTGTCAATCTGCACTCAACCGCATCGGATTACAGCTGTG 660
DB	1112 ATCAGTGTCCCTTTTGGCCCTGTCAATCTGCACTCAACCGCATCGGATTACAGCTGTG 1171
QY	661 CATGCTAAATACAGAGTCCCTGTGCTGGGAAACCAATTGATCAGTTGAAGTCAACCTGTG 720
DB	1172 CATGCTAAATACAGAGTCCCTGTGCTGGGAAACCAATTGATCAGTTGAAGTCAACCTGTG 1231
QY	721 CTGTGATATTTTCTGTTATTTGATGCTGGGGTGGAAATCCCATGGCAAGCTACTTCCAGAG 780
DB	1232 CTGTGATATTTTCTGTTATTTGATGCTGGGGTGGAAATCCCATGGCAAGCTACTTCCAGAG 1291
QY	781 GTTCCTCTCTTATCTCAGACCACTATAGCTCAGGTACTGCTTCCCTGGGACGCTTTTGGG 840
DB	1292 GTTCCTCTCTTATCTCAGACCACTATAGCTCAGGTACTGCTTCCCTGGGACGCTTTTGGG 1351
QY	841 TGCCCTGTGATGAGCTTACCCGCCATATGATAGAGCTATTTGAGGCTTCCACAGACTGG 900
DB	1352 TGCCCTGTGATGAGCTTACCCGCCATATGATAGAGCTATTTGAGGCTTCCACAGACTGG 1411
QY	901 AACCAAGACTGCTTACGGGTATCCAGATTCCAAGACTTAAGAGAGAACAGACATGATTTCTC 960
DB	1412 AACCAAGACTGCTTACGGGTATCCAGATTCCAAGACTTAAGAGAGAACAGACATGATTTCTC 1471



```

OY 961 CCGATCGTTCGACGATCCTCTGCCCCCTGTGTACATCTCTTTGGGGCTGGTGTCTT 1020
DB 1472 CCGATCGTTCGACGATCCTCTGCCCCCTGTGTACATCTCTTTGGGGCTGGTGTCTT 1531
OY 1021 TAGCTGCTGTATGTCCTTCAGCTGATGTCCTGTCCTCTGCGGAGTTCTATGTTGCT 1080
DB 1533 TAGCTGCTGTATGTCCTTCAGCTGATGTCCTGTCCTCTGCGGAGTTCTATGTTGCT 1591
OY 1081 CGGAATATCTACAGCTTCTTCAGCAAAATGCATCAGCAAGGAATGTGGGGTC 1140
DB 1592 CGGAATATCTACAGCTTCTTCAGCAAAATGCATCAGCAAGGAATGTGGGGTC 1651
OY 1141 ATAGAGATCACTGTGCTGTGTTCGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1200
DB 1652 ATAGAGATCACTGTGCTGTGTTCGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1711
OY 1201 ACTGTGTATGGGCTGTGTATCTGACCTGTGACCTTGTCTCATCATCTTCCACAG 1260
DB 1712 ACTGTGTATGGGCTGTGTATCTGACCTGTGACCTTGTCTCATCATCTTCCACAG 1771
OY 1261 CTGCTCTGTGTACTCTTCTCAAGAGAACACACTTATGGGGCAGTTGCTGTATAT 1320
DB 1772 CTGCTCTGTGTACTCTTCTCAAGAGAACACACTTATGGGGCAGTTGCTGTATAT 1831
OY 1321 TTTGACATATCTCTGAGAAATTAATCTGAGAGAGAGCCATATCTATCTTGAGCCCTTAATC 1380
DB 1832 TTTGACATATCTCTGAGAAATTAATCTGAGAGAGAGCCATATCTATCTTGAGCCCTTAATC 1891
OY 1381 TTTCTACCTGCTGTATTAATCTCTGACAGAAATGTATATACATCAGAGTTCCCATTTAA 1440
DB 1892 TTTCTACCTGCTGTATTAATCTCTGACAGAAATGTATATACATCAGAGTTCCCATTTAA 1951
OY 1441 ACTCTCTCATGATGTTACCTTCTTACCAATGTTGTTCTTATCAGCAAGTAT 1500
DB 1952 ACTCTCTCATGATGTTACCTTCTTACCAATGTTGTTCTTATCAGCAAGTAT 2011
OY 1501 CTATTTTGAAAGTGAACCTTGTGCTCCCAAAATTAGATGATTTTGTGCTGTGCGAAG 1560
DB 2012 CTATTTTGAAAGTGAACCTTGTGCTCCCAAAATTAGATGATTTTGTGCTGTGCGAAG 2071
OY 1561 CACAGTGAAGAGAACATGAGCAAGACCATTTCTAGTCAAGAAATGAATTAATTAAT 1620
DB 2072 CACAGTGAAGAGAACATGAGCAAGACCATTTCTAGTCAAGAAATGAATTAATTAAT 2131
OY 1621 GAACCTTGACCTGTGAACCTTGGGAGAGCCCTAACCTTCACTTCAACATTAAG 1680
DB 2132 GAACCTTGACCTGTGAACCTTGGGAGAGCCCTAACCTTCACTTCAACATTAAG 2191
OY 1681 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGGCTTGGAGTGAAGTAATTTACAA 1740
DB 2192 GAGGCCCTCTTGATGTTGATTCAGTCCGAGAGGGCTTGGAGTGAAGTAATTTACAA 2251
OY 1741 TGA 1743
DB 2252 TGA 2254

```

```

RESULT 2
AK034415 4306 bp mRNA linear HTC 02-SEP-2005
LOCUS AK034415
DEFINITION Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:931018K24 product:solute carrier family 5
(choline transporter), member 7, full insert sequence.
ACCESSION AK034415
VERSION AK034415.1 GI:26329926
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.

```

```

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
2 Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system -384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The PANTOM Consortium, the RIKEN Genome Exploration Research Group
Phase I and II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group
(Genome Network Core Team) and the PANTOM Consortium.
TITLE Antisense transcription in the Mammalian Transcriptome
JOURNAL Science 309, 1564-1566 (2005)
REFERENCE
AUTHORS The PANTOM Consortium, Riken Genome Exploration Research Group and
Genome Science Group (Genome Network Project Core Group).
TITLE The Transcriptional Landscape of the Mammalian Genome
JOURNAL Science 309, 1559-1563 (2005)
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, D., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail: genome-res@gscl.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES Location/Qualifiers

```

```
source
1. 4306
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330188K24"
/db_xref="taxon:10090"
/clone="9330188K24"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
394_2136
/note="unnamed protein product; putative
solute carrier family 5 (choline transporter), member 7
MGDI[MGI:1927126 GB|Nm_022025, evidence: BLASTN, 99%,
match=1743]"
/codon_start=1
/protein_id="BAC28702.1"
/db_xref="GI:26329927"
/translation="MSFVHESGLVAIIIFYLLIFLVGIWAMKTKNSGNEPSEAIIV
GGRDGLVGGFTMTAVWGGYINGTAEAVYRPGCLAMAPAGYSLILGLF
AKPMRSKGYTMLDPFOIYGRMGGLFIIPALMEEMFAAIFPALGATTSVIDV
VNISIVSALIAIIVTIGLVSVAYTDVOLFCEIPIGLMISVPSALSHPATVDIGPT
AVHAKYOSPMWLTIESVEVYTMIDNFIILMGIPIWQAYFQPIVSSSATYVPSFL
AAFGCLVMAIPAICIGAIASIDMNOTAYGFPDPTKEADTILYILOYICPVYISF
FGLGAVSAVMSADSSILSASMFARINYSFRONSDEKIVWMRTIVLVFASATA
TAMALLTKTVGLMWSLDVYIIIPOLLVLFKGNNTGAVAGYIFGLIYRTGG
EPYILOPLIFPYGYSDKNGIYNORFPEKLSMTVSFTFNI CVSYIAKYLFESGLP
PKLDVEAVVARHSEEMDKITLVRENENIKINELAPVKPRQSLTISFTFNKALLDV
DSPSGSGTEBDLQ"
```

Query Match 99.4%; Score 1731.8; DB 6; Length 4306;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1736; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
1 ATGCTTTTCATGTGGAGAGACTGTGATGCTATATTCCTCTTCTACCTCCTTATTTTCG 60
394 ATGTCTTTCCAGTGAAGAGACTGTGATGCTATATTCCTCTTCTACCTCCTTATTTTCG 453
61 GTTGGAAATATGGGCTGCATGAAAAACCAAAAACAGCGGCAACCCAGAAAGCGCAGTGA 120
454 GTTGAATATATGGGCTGCATGAAAAACCAAAAACAGCGGCAACCCAGAAAGCGCAGTGA 513
121 GCCATCATATGTCGGGCGCGTGAATGTTGTTGTTGTTGTTTTCATGACAGCC 180
514 GCCATCATATGTCGGGCGCGTGAATGTTGTTGTTGTTGTTTTCATGACAGCC 573
181 ACCGTGGTTGGAGAGGCTACATCAATGGGAGACAGACAGCTGATGGGCGAGTTGT 240
574 ACCGTGGTTGGAGAGGCTACATCAATGGGAGACAGACAGCTGATGGGCGAGTTGT 633
241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 300
634 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCTAATTTTAAAGTGTCTG 693
301 TTTTTCGCAAACTATATGCTTCAAGGGATATGTACTATGTATGTAACCATTTCAACAG 360
694 TTTTTCGCAAACTATATGCTTCAAGGGATATGTACTATGTATGTAACCATTTCAACAG 753
361 ATCTATGAAAGCGATGGGTGGGCTGCTTATCCCGTCACTGATGGGAGAGATGTC 420
754 ATCTATGAAAGCGATGGGTGGGCTGCTTATCCCGTCACTGATGGGAGAGATGTC 813
421 TGGGCTGACGCAATTTCTCTGATTAAGGGCCACCATCAGCGTGAATGATGATGAT 480
814 TGGGCTGACGCAATTTCTCTGATTAAGGGCCACCATCAGCGTGAATGATGATGATGAT 873
481 GTGAACATATCGGTCAATGTTCTCTGACATCATTCGATTTCTTATACCTTAACTAGTGGTGG 540
874 GTGAACATATCGGTCAATGTTCTCTGACATCATTCGATTTCTTATACCTTAACTAGTGGTGG 933
541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACCTGTGG 600
```

```
Db
934 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGACCTGTGG 993
601 ATCAGTGTCCCTTTTGGCCCTGCTCAATCCTGACAGTCAACCGACATGCGATTCGAGCTGTG 660
994 ATCAGTGTCCCTTTTGGCCCTGCTCAATCCTGACAGTCAACCGACATGCGATTCGAGCTGTG 1053
661 CATGCTAAATACAGAGTCCCTGCTGGGAACCATTTGAATCATGTTGAAGTCTACACCTGG 720
1054 CATGCTAAATACAGAGTCCCTGCTGGGAACCATTTGAATCATGTTGAAGTCTACACCTGG 1113
721 CTGTATATTTTCTGTATATGATGCTGGGTGGAATCCCATGGCAACCTACTTCCAGAGG 780
1114 CTGTATATTTTCTGTATATGATGCTGGGTGGAATCCCATGGCAACCTACTTCCAGAGG 1173
781 GTCTCTCTTCAATCCTGACCAACCTATGCTGAGTACTGTCTTCTGCGACGCTTTTGGG 840
1174 GTCTCTCTTCAATCCTGACCAACCTATGCTGAGTACTGTCTTCTGCGACGCTTTTGGG 1233
841 TCCCTGTGATGCTCTTACCGGCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 900
1234 TCCCTGTGATGCTCTTACCGGCATATGATAGAGCTATTTGAGCTTCCACAGACTGG 1293
901 AACGACATGCTTACGGGTATCCAGATCCCAAGACTAAGAGGAGAGACAGACATGATTC 960
1294 AACGACATGCTTACGGGTATCCAGATCCCAAGACTAAGAGGAGAGACAGACATGATTC 1353
961 CCGATGCTTGCAGTACCTGCTGCGCTGTAATCATCTCTTCTTGGGCTGTGCTGTT 1020
1354 CCGATGCTTGCAGTACCTGCTGCGCTGTATACATCTCTTCTTGGGCTGTGCTGTT 1413
1021 TCAGCTGCTGATGCTCTTACGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
1414 TCAGCTGCTGATGCTCTTACGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
1081 CGGAATATCTACCACTTCCCTTCCAGCAAAATGATGATGACAGCAAGAAATGTGGGCTC 1140
1474 CGGAATATCTACCACTTCCCTTCCAGCAAAATGATGATGACAGCAAGAAATGTGGGCTC 1533
1141 ATGAGATCATCTGCTGCTTGTTCGAGCATCTGCAAGACCATGCTTGTGCTGACGAAG 1200
1534 ATGAGATCATCTGCTGCTTGTTCGAGCATCTGCAAGACCATGCTTGTGCTGACGAAG 1593
1201 ACTGTGATGGGCTCTGGTACTGAGCTCTGACCTTGTCTATCATATCTTCCACAG 1260
1594 ACTGTGATGGGCTCTGGTACTGAGCTCTGACCTTGTCTATCATATCTTCCACAG 1653
1261 CTGCTCTGTATCTTTCATCAAAAGAACCAACATTTATGGGCGAGTTGCTGTTATAT 1320
1654 CTGCTCTGTATCTTTCATCAAAAGAACCAACATTTATGGGCGAGTTGCTGTTATAT 1713
1321 TTTGACTATTTCTGAGATTAATGAGAGAGACCATATCTATCTTGACGCCCTTAATC 1380
1714 TTTGACTATTTCTGAGATTAATGAGAGAGACCATATCTATCTTGACGCCCTTAATC 1773
1381 TTCTAACCTGTTATTAATCTGACAAAGATGTATATACATCAGAGTTCCCATTTAA 1440
1774 TTCTAACCTGTTATTAATCTGACAAAGATGTATATACATCAGAGTTCCCATTTAA 1833
1441 ACTCTCTGCATGCTACCTCATTTCTTACCAACATTTGTTGTTCTTATCTAGCAAGTAT 1500
1834 ACTCTCTGCATGCTACCTCATTTCTTACCAACATTTGTTGTTCTTATCTAGCAAGTAT 1893
1501 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGATGATGCTGTGTGCAAGG 1560
1894 CTATTTGAAAGTGAACCTTGGCTCCAAATTTAGATGATGATGCTGTGTGCAAGG 1953
1561 CACAGTGAAGAGAAATGAGCAAGACATTTCTAGTCAAGAAATGAAATATCAATTTAAT 1620
1954 CACAGTGAAGAGAAATGAGCAAGACATTTCTAGTCAAGAAATGAAATATCAATTTAAT 2013
1621 GAACCTGCACTGTGAAACCTGCGAGAGCTTAACCTCAGTTCAACTTCCACATAAG 1680
```

Db 2014 GAACCTGCACCTGTGAACCTCGGAGAGCCTAACCTCAGTTCAACTTTCACCAATAAG 2073

QY 1681 GAGGCCCTCGTATGTTGATTCAGTCCGAGAGGGCTGGGACTGAAGTAATTTAGCA 1740

Db 2074 GAGGCCCTCGTATGTTGATTCAGTCCGAGAGGGCTGGGACTGAAGTAATTTAGCA 2133

QY 1741 TGA 1743

Db 2134 TGA 2136

RESULT 3  
AY413300 1743 bp DNA linear GSS 12-DEC-2003  
LOCUS Mus musculus HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY413300  
VERSION AY413300.1 GI:39769262  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1743)  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source location/Qualifiers  
1..1743  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>1743  
/locus\_tag="HCM4844"

ORIGIN

Query Match 84.1%; Score 1465; DB 14; Length 1743;  
Best Local Similarity 84.2%; Pired. No. 0;  
Matches 1468; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1 ANGCCCTTCATGGAAGAGCTGTAGCTATATCTCTTCACTCCCTTATATTTCTG 60

Db 1 ATGTCTTTCCAGCTAGAGAGCTGTAGCTATATCTCTTCACTCCCTTATATTTCTG 60

QY 61 GTTGGAATATGGGCTGCATGGAAGAACCAAAAACAGGGGCAACCCAGAGAGCGCAGTAA 120

Db 61 GTTGGAATATGGGCTGCATGGAAGAACCAAAAACAGGGGCAACCCAGAGAGCGCAGTAA 120

QY 121 GCATCATATAGTGGGGGGCTGCATTTGTTGTTGTTGTTGTTTACATGACAGCC 180

Db 121 GCATCATATAGTGGGGGGCTGCATTTGTTGTTGTTGTTTACATGACAGCANN 180

QY 181 ACCTGGGTTGAGAGAGCTCATCATATGAGACAGACAGAGTGTATGGCCAGGTTGT 240

Db 181 NNN 240

QY 241 GGTCTAGCTTGGGCTCATGACCACTTGATTTCTCTGAGCTAATTTTAAAGTGCTG 300

Db 241 NNN 300

QY 301 TTTTTCGCAACCTATGGCTTCCAGGAGATATGTACTATGTTAGACCATTCACAACG 360

Db 301 NNN 360

QY 361 ATCTATGAAAGGAGCTGGGTGGGCTGCTTTCATCCCTGCACCTGATGGAGAGATGTTT 420

Db 361 NNN 420

QY 421 TGGGCTGCAGCAATTTCTCTGATTAAGGGGCCACCATCAGCTGATCATGATGAGAT 480

Db 421 NNN 480

QY 481 GTGAACATATCGGTCAATGTTCTGCACTCATTTGCCATTTATACCTAGTGGGTGG 540

Db 481 GTGAACATATCGGTCAATGTTCTGCACTCATTTGCCATTTATACCTAGTGGGTGG 540

QY 541 CTCTACTGTGGCATATATCTGATGTTGTCAGCTATTCGCAATTTTATAGGACTGTGG 600

Db 541 CTCTACTGTGGCATATATCTGATGTTGTCAGCTATTCGCAATTTTATAGGACTGTGG 600

QY 601 ATCAGTGTCCCTTTTCCCTGTCACATCTGCAGTCAACGACATCGGATTCACAGCTGTG 660

Db 601 ATCAGTGTCCCTTTTCCCTGTCACATCTGCAGTCAACGACATCGGATTCACAGCTGTG 660

QY 661 CATGCTAATATCCAGAGTCCCTGGCTGGGAGAACATGGAATCAATGAAGCTACACTGG 720

Db 661 CATGCTAATATCCAGAGTCCCTGGCTGGGAGAACATGGAATCAATGAAGCTACACTGG 720

QY 721 CTGTATATATTTCTGTATTTGATGCTGGGTGAATCCCATGGCAAGCTTCTCCAGAG 780

Db 721 CTGTATATATTTCTGTATTTGATGCTGGGTGAATCCCATGGCAAGCTTCTCCAGAG 780

QY 781 GTCTCTCTTCATCTCTGACGACCATATGCTCAGATAGTGTCTTCTGCGAGCTTTGGG 840

Db 781 GTCTCTCTTCATCTCTGACGACCATATGCTCAGATAGTGTGTCTTCTGCGAGCTTTGGG 840

QY 841 TGGCTGTGATGGCTTACCCGACATATGATGAGAGCTATGAGAGTTCCACAGACTGG 900

Db 841 TGGCTGTGATGGCTTACCCGACATATGATGAGAGCTATGAGAGTTCCACAGACTGG 900

QY 901 AACCAAGTCCCTCAACGGATATCCAGATCCCAAGACTAAGAGAGACAGACATGATTTCT 960

Db 901 AACCAAGTCCCTCAACGGATATCCAGATCCCAAGACTAAGAGAGACAGACATGATTTCT 960

QY 961 CCGATGTTCTGAGTACCTGTGCTGTGTATCATCTCTTCTTTGGGCTTGGTCTGTT 1020

Db 961 CCGATGTTCTGAGTACCTGTGCTGTGTATCATCTCTTCTTTGGGCTTGGTCTGTT 1020

QY 1021 TCACTGCTGTATGTCATGTCCTGACGCTGCTGATCTGTGCGGCGATTTAGTTGCT 1080

Db 1021 TCACTGCTGTATGTCATGTCCTGACGCTGCTGATCTGTGCGGCGATTTAGTTGCT 1080

QY 1081 CGGAATATCTACAGCTTTCTTCCAGCAAAATGATCAGACAGAAATGTTGTGGTCT 1140

Db 1081 CGGAATATCTACAGCTTTCTTCCAGCAAAATGATCAGACAGAAATGTTGTGGTCT 1140

QY 1141 ATGAGGATCACTGTGCTTGTGTGAGAGATCTGCAACAGCCATGCTTGTGTCAGAG 1200

Db 1141 ATGAGGATCACTGTGCTTGTGTGAGAGATCTGCAACAGCCATGCTTGTGTCAGAG 1200

QY 1201 ACTGTGTATGGGCTGTGATCTGACCTGCACTTGTCTATCATCATCTTCCACAG 1260

Db 1201 ACTGTGTATGGGCTGTGATCTGACCTGCACTTGTCTATCATCATCTTCCACAG 1260

QY 1261 CTGCTGTGTATCTTCTTCAATCAAGAGAACCAACATTTATGGGCAAGTGTGTTATTT 1320

Db 1261 CTGCTGTGTATCTTCTTCAATCAAGAGAACCAACATTTATGGGCAAGTGTGTTATTT 1320

QY 1321 TTTGACATATCTGTGAATTTACTGAGAGAGACCATATCTAATCTGAGCCCTTAATC 1380

Db	1321	TTTGGACATTTCTCGGAATTA	1743 bp	DNA	linear	GSS 12-DEC-2003	1380
Qy	1381	TTTCACCTCGTTATTACTCTGACAGAA					1440
Db	1381	TTTCACCTCGTTATTACTCTGACAGAA					1440
Qy	1441	ACTCTCCATGAGTAACTCATTTCTTACCA					1500
Db	1441	ACTCTCTCATGATTAACCTCATTTCTTACCA					1500
Qy	1501	CTATTTGAAAGTGAACCTTGCCCTCCAA					1560
Db	1501	CTATTTGAAAGTGAACCTTGCCCTCCAA					1560
Qy	1561	CACACTGGAAGAAATGAGACAAACCA					1620
Db	1561	CACAGTGAAGAAATGAGACAAACCA					1620
Qy	1621	GAACTTGACCTGTGAAACCTGCGAGAG					1680
Db	1621	GAACTTGACCTGTGAAACCTGCGAGAG					1680
Qy	1681	GAGGCCCTCTTGATGTTGATTCACGT					1740
Db	1681	GAGGCCCTCTTGATGTTGATTCACGT					1740
Qy	1741	TGA 1743					
Db	1741	TGA 1743					
RESULT 4	AY413298	1743 bp	DNA	linear	GSS 12-DEC-2003		
LOCUS	AY413298	1743 bp	DNA	linear	GSS 12-DEC-2003		
DEFINITION	AY413298	1743 bp	DNA	linear	GSS 12-DEC-2003		
ACCESSION	AY413298	1743 bp	DNA	linear	GSS 12-DEC-2003		
VERSION	AY413298.1	GI:39769260					
KEYWORDS	GSS.						
SOURCE	Homio sapiens (human)						
ORGANISM	Homio sapiens						
REFERENCE	1 (bases 1 to 1743)						
AUTHORS	Homio sapiens						
TITLE	Homio sapiens						
JOURNAL	Science 302 (5652), 1960-1963 (2003)						
PUBLISHED	14671302						
REFERENCE	2 (bases 1 to 1743)						
AUTHORS	Homio sapiens						
TITLE	Homio sapiens						
JOURNAL	Science 302 (5652), 1960-1963 (2003)						
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment						
FEATURES	Location/Qualifiers						
source	1. 1743						
gene	/mol_type="genomic DNA"						
ORIGIN	<1..>1743						
Query Match	66.7%; Score 1162.6; DB 14; Length 1743;						

Db 1021 TCTGCTGCTGTTATGTCATCAGCAGATTCCTTCATCTTGTCAGCAAGTTCCATGTTTGA 1080  
 Oy 1081 CGGAATATCTACAGAGCTTCTCCAGACAAATGATCAGACAGAGAAATGTTGTTGTC 1140  
 Db 1081 CGGAATATCTACAGAGCTTCTCCAGACAAATGATCAGACAGAGAAATGTTGTTGTC 1140  
 Oy 1141 ATGAGATCATCTGCTGTTGTTGAGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1200  
 Db 1141 ATGAGATCATCTGCTGTTGTTGAGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1200  
 Oy 1201 ACTGTGTATGAGGCTCTGTTGTTGAGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1260  
 Db 1201 ACTGTGTATGAGGCTCTGTTGTTGAGAGATCTGCAACAGCCATGCTTGTGTCAGGAAG 1260  
 Oy 1261 CTGCTCTGTGTTACTCTTCAACAAAGAGACCAACATTAAGGGGAGTTGCTGTTATTT 1320  
 Db 1261 CTGCTCTGTGTTACTCTTCAACAAAGAGACCAACATTAAGGGGAGTTGCTGTTATTT 1320  
 Oy 1321 TTTGACATATCTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Db 1321 TTTGACATATCTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Oy 1381 TTTGACATATCTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
 Db 1381 TTTGACATATCTCTGAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
 Oy 1441 ACTCTCTGAG 1500  
 Db 1441 ACTCTCTGAG 1500  
 Oy 1501 CTATTTGAG 1560  
 Db 1501 CTATTTGAG 1560  
 Oy 1561 CACAGTGAAG 1620  
 Db 1561 CACAGTGAAG 1620  
 Oy 1621 GAACCTGAG 1680  
 Db 1621 GAACCTGAG 1680  
 Oy 1681 GAGGCTCTCTGATGTTGATTCAGTCCGAGAGGCTGAGAGTGAAGTAAATTACAA 1740  
 Db 1681 GAGGCTCTCTGATGTTGATTCAGTCCGAGAGGCTGAGAGTGAAGTAAATTACAA 1740  
 Oy 1741 TGA 1743  
 Db 1741 TGA 1743

RESULT 5  
 AY413299 1743 bp DNA linear GSS 12-DEC-2003  
 LOCUS Pan troglodytes HCM4844 gene, VIRUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION AY413299  
 VERSION AY413299.1 GI:39769261  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Pan.  
 1 (bases 1 to 1743)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,  
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302

REFERENCE 2 (bases 1 to 1743)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,  
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment  
 FEATURES  
 Location/Qualifiers  
 source 1..1743  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 <1..>1743  
 /locus\_tag="HCM4844"  
 ORIGIN  
 Query Match 66.7%; Score 1162.2; DB 14; Length 1743;  
 Best Local Similarity 73.3%; Pred. No. 0;  
 Matches 1278; Conservative 0; Mismatches 465; Indels 0; Gaps 0;  
 Oy 1 ATGCCCTTTCATGTGAG 60  
 Db 1 ATGCCCTTTCATGTGAG 60  
 Oy 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGGCGCAACCCAAAGAGCGAGATGAA 120  
 Db 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGGCGCAACCCAAAGAGCGAGATGAA 120  
 Oy 121 GCCATCATAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180  
 Db 121 GCCATCATAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180  
 Oy 181 ACTGTGTTGAG 240  
 Db 181 ACTGTGTTGAG 240  
 Oy 241 GGCTGAGCTTGGGCTCATGACCCATGATATCTCTGAGTCAATTTAGTGGTCTG 300  
 Db 241 GGCTGAGCTTGGGCTCATGACCCATGATATCTCTGAGTCAATTTAGTGGTCTG 300  
 Oy 301 TTTTTCGCAACCTATGCTTCCAGAGATATGTACTATGTAGACCATCAAAACAG 360  
 Db 301 TTTTTCGCAACCTATGCTTCCAGAGATATGTACTATGTAGACCATCAAAACAG 360  
 Oy 361 ATCTATGAAAGCGCATGGGTGGGCTCTTCATCCTGCACTGATGGAGAGATGTTTC 420  
 Db 361 ATCTATGAAAGCGCATGGGTGGGCTCTTCATCCTGCACTGATGGAGAGATGTTTC 420  
 Oy 421 TGGGCTGCAAGATTTTCTCTGCAATTAAGGGCCACATCAGGCTATCATATGTTGAT 480  
 Db 421 TGGGCTGCAAGATTTTCTCTGCAATTAAGGGCCACATCAGGCTATCATATGTTGAT 480  
 Oy 481 GTGAACATATCGGCTCATGCTCTGCACTCATGTCATTTTATACCTAGTGGTGGG 540  
 Db 481 GTGAACATATCGGCTCATGCTCTGCACTCATGTCATTTTATACCTAGTGGTGGG 540  
 Oy 481 ATGCACATTTCTGTATCATCTCTGCACTCATGTCATTTTATACCTAGTGGTGGG 540  
 Db 481 ATGCACATTTCTGTATCATCTCTGCACTCATGTCATTTTATACCTAGTGGTGGG 540  
 Oy 541 CTCTACTCTGTGGCATATATGATGTTGTCAGACTATTCATGCAATTTTATAGAGCTGTG 600  
 Db 541 CTCTACTCTGTGGCATATATGATGTTGTCAGACTATTCATGCAATTTTATAGAGCTGTG 600  
 Oy 541 CTCTATTTGTGGCTTACACTGATGTGCTTTCAGCTCTTTTGTAGAGGCTGTGG 600  
 Db 541 CTCTATTTGTGGCTTACACTGATGTGCTTTCAGCTCTTTTGTAGAGGCTGTGG 600  
 Oy 601 ATCAGTGTCCCTTTGCTGATGTCATCTGCACTGCAAGTGCAGACATCGGCTTCACTGCTGTG 660  
 Db 601 ATCAGTGTCCCTTTGCTGATGTCATCTGCACTGCAAGTGCAGACATCGGCTTCACTGCTGTG 660  
 Oy 661 CATGCTAAATACAGAGTCCCTGCTGAGAACCATTTGAATCACTTGAAGTCAACCTGG 720  
 Db 661 CATGCTAAATACAGAGTCCCTGCTGAGAACCATTTGAATCACTTGAAGTCAACCTGG 720  
 Oy 721 CTGTGAATTTTCTGTATTTGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780  
 Db 721 CTGTGAATTTTCTGTATTTGATGCTGGGTGGAATCCATGGCAAGCTTCCAGAGG 780

Db	72	CTTAAATAGTTCCTTGTTGTATGATGCTGGGGTGAATTCACATGGCAAGCAACTTTCAAGG	780
QY	781	GTCTCTCTTCAATCCTCAAGCCACTAAGCTCAAGGTATCTGTCTCTTCTGGCAGCTTTTGG	840
Db	781	GTTCCTCTCTTCTCTTCTGAGCCACTAATGCTCAAGAGCTGTCTCTTCTGGCAGCTTTGCGG	840
QY	841	TGCGTGTGATGAGCTCTACCCGCCATATGACATAGAGAGCTAATTGGAGCTTCCACAGACTGG	900
Db	841	TGCCTGTGATAGGCCATCCAGACCACTATCATTCGGGGCCATTGGAGATCAACAGACTGG	900
QY	901	AACCAAGATCGCCCTACGGGTATCCAGATCCCAAGACTAAGAGGAAAGCAGACATGATTC	960
Db	901	AACCAAGATCGATATGGGCTTCCAGATCCCAAGACTAAGAGGAGGCGAGATATTTTA	960
QY	961	CCGATCGTTCTGACAGTACTCTGCCCTGTGATACATCTCTCTTTGGGCTTGGTGGT	1020
Db	961	CCATTTGTTCTGACAGTACTCTGCCCTGTGATATTTCTTTCTTTGGCTTGGTGGAGTT	1020
QY	1021	TCAGCTGCTGATGCTGCTCAAGCTGACTGTCATCTGTGCGGCGAGTTTATGTTGCT	1080
Db	1021	TCGTGCTGCTGATATGCTATCAAGACAGATTTCTTCATCTGTGAGGAGATTTCCATGTTTCA	1080
QY	1081	CGGAATATCTACAGACTTTCTTCAAGCAAAAATGATCAAGCAAGAAATGTGTGGGTC	1140
Db	1081	CGGAACATCTACAGACTTTCTTCAAGCAAAAATGTTGGCAAAAGATGTTTGGGTT	1140
QY	1141	ATGAGATACACTGTGCTGTGTTCGAGACATCTGACAGCCATGGCTTGTCTGACGAG	1200
Db	1141	ATGCGAATCAAGTGTTGTGTGTGGACATCTGACAGCCATGGCTTGTCTGACGAG	1200
QY	1201	ACTGTGTATGGGCTTGTGATCTGAGCTCTGACTGTGTCTATCATCATATCTTCCACAG	1260
Db	1201	ACTGTGTATGGGCTCTGTGATCTGACTGATCTGACTGTGTCTATCATCTTCCCCACG	1260
QY	1261	CTGCTCTGTGACTCTTTCATCAAGAGAACAACTTATGGGGGAGTGTGGTGTATAT	1320
Db	1261	CTGCTTGTGTACTCTTGTGTATAGGAAACAACTTATGGGGCCGTGGAGTTATGTT	1320
QY	1321	TTTGGACTATTCTGAGATTAAGTGAGAGAGACCATATCTATACTTGACGCCCTTATC	1380
Db	1321	TCGTGCCCTCTCTGAGATTAAGTGAGAGGAGCCATATCTGTATCTTCAACCCCTTGATC	1380
QY	1381	TTTCAACCTGGTTATTACTCTGACAAAGATGTATATCAATCAAGAGTTCCTTAA	1440
Db	1381	TTTCAACCTGGCTATTAACCTGATGATATGATGTATATCAAGAAATTCATTTAA	1440
QY	1441	ACTCTCTCATGTATACCTCAATCTTTTACCAACATTTGTGTTCTTATCTACGCAAGTAT	1500
Db	1441	ACACTGGCCATGGTTAAGCATTTCTTAAACCAACATTTGCTCTCTATCTAGCCAAATAT	1500
QY	1501	CTATTTTGAAAGTGGAACTTGTGCTTCAAAATTAAGTATGATTTGATGCTGTTGTGGCAAG	1560
Db	1501	CTATTTTGAAAGTGGAACTTGTGCAACTTAAATTAAGTATGATTTGATGCTGTTGTGCAAGA	1560
QY	1561	CACAGTGAAGAGACATGGACAGACCACTTCTAGTCACAAAATGAAAATATCAAAATTAAT	1620
Db	1561	CACAGTGAAGAAACATGGATGATGACATTTCTTGTCAAAAATGAAAATATTAATTTAGAT	1620
QY	1621	GAACTTGCACCTGTGAAACCGACAGAGAGCATGAGCCCTCAGCTCAACTTCAACCAATTA	1680
Db	1621	GAGGCGCCCTCTGATGATGATTCAGAGTCGAGAGGGGCTGGGAGCTGAAGATTAATTTCAA	1740
QY	1681	GAGGCGCTTCTTGATGTGATTTCACTGACAGANAGCTGGGACTGAAGATTAATTTACAG	1740
Db	1741	TGA 1743	
	1741	TGA 1743	
	1741	TGA 1743	

LOCUS	AK164116	4122 bp	RNA	linear	HTC 21-SEP-2005	
DEFINITION	Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone:CS30033506 product:solute carrier family 5 (choline transporter), member 7, full insert sequence.					
ACCESSION	AK164116					
VERSION	AK164116.1	GI:74211080				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus					
ORGANISM	Mus musculus (house mouse)					
REFERENCE	Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Euarchontoglires, Glires, Rodentia, Surognath, Muroidae, Muridae, Murinae, Mus.					
AUTHORS	1	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning					
JOURNAL	2	Ichih, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
PUBMED	10349636	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
REFERENCE	Genome Res. 10 (10), 1617-1630 (2000)					
JOURNAL	11042159					
PUBMED	11042159					
AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, K., Nishi, K., Kitanishi, T., Tashiro, H., Harada, A., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format					
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)					
PUBMED	11076861					
REFERENCE	4	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Komoto, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kado, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hune, D. A., Kamiya, M., Lee, N.H., Lyons, P., Machioni, L., Mashima, J., Mazzarelli, S., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokawa, K., Wang, K.H., Weitz, C., Wiltaker, C., Wilming, L., Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.				
CONSTRM	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium					
TITLE	Functional annotation of a full-length mouse cDNA collection					
JOURNAL	Nature 409 (6821), 685-690 (2001)					
PUBMED	11217851					
REFERENCE	5	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Ose, N., Saito, R., Suzuki, H., Yamataka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Kyochawa, C., Gotohori, T., Baldarelli, R., Hill, D., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L.M., Kanpin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brack, D., Brusic, V., Chottha, C., Corbani, L.E., Cousins, S., Dalla, E., Dresani, T.A., Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,				





Db 486 TCTATGTTGCTCGGAATATCTACAGCTTCTCCTCAGACAAATGCATCAGACAGGA 545  
 QY 1129 ATTGTGGGTGATGAGATCACTGTGCTTGTGTGAGACATCTGACACACCATGCT 1188  
 Db 546 ATGTGTGGGTATGAGATCACTGTGCTTGTGTGAGACATCTGACACACCATGCT 605  
 QY 1189 TTGCTGACGAGACATGATGATGGGCTCTGGTACCTGAGCTTGTCTACATCATC 1248  
 Db 606 TTGCTGACGAGACATGATGATGGGCTCTGGTACCTGAGCTTGTCTACATCATC 665  
 QY 1249 ATCTCCACAGCTGCTGTGTACTCTTCATCAAGAACCAACCTATGAGGAGGT 1308  
 Db 666 ATCTCCACAGCTGCTGTGTACTCTTCATCAAGAACCAACCTATGAGGAGGT 725  
 QY 1309 GCTGTTATATTTTGTGACTATCTGAGAAATTAATGAGAGACCATATATATCTTG 1368  
 Db 726 GCTGTTATATTTTGTGACTATCTGAGAAATTAATGAGAGACCATATATATCTTG 785  
 QY 1369 CAGCCTTATATTTTGTGACTATCTGAGAAATTAATGAGAGACCATATATATCTTG 1428  
 Db 786 CAGCCTTATATTTTGTGACTATCTGAGAAATTAATGAGAGACCATATATATCTTG 845  
 QY 1429 TTCCATTTTAAATCTCTCCATGTTACCTGTTACCAACATTTGTTCTTAT 1488  
 Db 846 TTCCATTTTAAATCTCTCCATGTTACCTGTTACCAACATTTGTTCTTAT 905  
 QY 1489 CTAGCAGATATCTATTTGAAAGTGAACCTTGCTCCAAATTTAGATGATTTGATGCT 1548  
 Db 906 CTAGCAGATATCTATTTGAAAGTGAACCTTGCTCCAAATTTAGATGATTTGATGCT 965  
 QY 1549 GTTGTGCAAGGCAAGTGAAGAACTGAGACAGACCATTTCTAGTCGAATATGAAT 1608  
 Db 966 GTTGTGCAAGGCAAGTGAAGAACTGAGACAGACCATTTCTAGTCGAATATGAAT 1025  
 QY 1609 ATCAATTTAAATGAATCTTGACCTGTGAAACCTCGGACAGACTTAACCTCGATTTCACT 1668  
 Db 1026 ATCAATTTAAATGAATCTTGACCTGTGAAACCTCGGACAGACTTAACCTCGATTTCACT 1085  
 QY 1669 TTACCAATTAAGAGAGCCCTCTTGTATGTTGATTCAGTCCGAGAGGCTCTGGACTGAA 1728  
 Db 1086 TTACCAATTAAGAGAGCCCTCTTGTATGTTGATTCAGTCCGAGAGGCTCTGGACTGAA 1145  
 QY 1729 GATATTTTACATGA 1743  
 Db 1146 GATTAACCTTACATGA 1160

RESULT 7  
 CD350164 707 bp mRNA linear EST 09-JUL-2003  
 DEFINITION U1-M-FY0-cfl-h-10-0-UI-r1 NIH-BMAP\_FY0 Mus musculus cDNA clone  
 IMAGE:6851099 5', mRNA sequence.  
 ACCESSION CD350164  
 VERSION CD350164.1 GI:31141679  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.  
 1 (bases 1 to 707)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rmail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/mousefl.html  
 This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)  
 Seq primer: pyx-5.  
 Location/Qualifiers  
 1..707  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6851099"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"  
 /lab\_host="MD10B (T1 phage resistant)"  
 /clone\_lib="NIH-BMAP\_FY0"  
 /note="Organ: Brain; Vector: pyx-Asc; Site: 1: EcoR I; Site: 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

Query Match 34.6%; Score 603.4; DB 5; Length 707;  
 Best Local Similarity 98.4%; Pred. No. 1.8e-169;  
 Matches 662; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

1 ATGCTTTCCATGATGAGAGAGCTGATCTATATCTCTTCTACCTCTTATATTTCTG 60  
 27 ATGTCTTTCCAGTGAAGAGAGCTGATCTATATCTCTTCTACCTCTTATATTTCTG 86  
 61 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGGCAACCCAGAAAGCGCAGTGA 120  
 87 GTTGAATATGAGGCTGATGAGAAACCAAAACAGCGGCAACCCAGAAAGCGCAGTGA 146  
 121 GGCATCATAGTCGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTACCATGACAGC 180  
 147 GGCATCATAGTCGGGGGCGGTGACATTTGTTGTTGTTGTTGTTTACCATGACAGC 206  
 181 ACCTGGTTGAGAGAGAGCTATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGT 240  
 207 ACCTGGTTGAGAGAGAGCTATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGT 266  
 241 GGTCTAGCTTGGGCTCATGACACCATTTGATATTTCTGAGCTTATATTTTATGAGTCTG 300  
 267 GGTCTAGCTTGGGCTCATGACACCATTTGATATTTCTGAGCTTATATTTTATGAGTCTG 326  
 301 TTTTGTGGAACCAATGCGTTCCAGGAGATATGATGATGATGATGATGATGATGATGATG 360  
 327 TTTTGTGGAACCAATGCGTTCCAGGAGATATGATGATGATGATGATGATGATGATGATG 386  
 361 ATCTATGGAAGCGATGAGGAGGCTGCTTTCATCTGACCTGACCTGATGGG-AGAGATGT 419  
 387 ATCTATGGAAGCGATGAGGAGGCTGCTTTCATCTGACCTGACCTGATGGAGAGATGT 446  
 420 CTGGGCTCAGCAATTTTCTGCTATGAGGAGGACCA-TCAGCGTATCATT-GATGTG 477  
 447 CTGGGCTCAGCAATTTTCTGCTATGAGGAGGACCA-TCAGCGTATCATT-GATGTG 506  
 478 GATGGAACATATGCGTATGATGCTCTGCACTCATGACATTTCTTATACCTATAGGCT 537  
 507 GATGGAACATATGCGTATGATGCTCTGCACTCATGACATTTCTTATACCTATAGGCT 566  
 538 GGGCTCTACTCTGAGCATATAT- GATGTTGTCAGGCTATTTGATTTTATAGACT 596  
 567 GGGCTCTACTCTGAGCATATAT- GATGTTGTCAGGCTATTTGATTTTATAGACT 626



RESULT 8	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
597	GT -GGATAGAGTCCCTTTTCGCCGTGACATCCGTGACGATCAGCAGATTCGATTCACAG 655	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
627	GTGGGATAGAGTCCCTTTTCGCCGTGACATCCGTGACGATCAGCAGATTCGATTCACAG 686	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
656	CTGTGCATGCTTAA 668	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
687	CTGTGCATGCTTAA 699	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
699	CTGTGCATGCTTAA 710	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
710	CTGTGCATGCTTAA 721	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
721	CTGTGCATGCTTAA 732	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
732	CTGTGCATGCTTAA 743	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
743	CTGTGCATGCTTAA 754	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
754	CTGTGCATGCTTAA 765	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
765	CTGTGCATGCTTAA 776	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
776	CTGTGCATGCTTAA 787	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
787	CTGTGCATGCTTAA 798	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
798	CTGTGCATGCTTAA 809	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
809	CTGTGCATGCTTAA 820	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
820	CTGTGCATGCTTAA 831	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
831	CTGTGCATGCTTAA 842	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
842	CTGTGCATGCTTAA 853	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
853	CTGTGCATGCTTAA 864	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
864	CTGTGCATGCTTAA 875	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
875	CTGTGCATGCTTAA 886	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
886	CTGTGCATGCTTAA 897	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
897	CTGTGCATGCTTAA 908	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
908	CTGTGCATGCTTAA 919	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
919	CTGTGCATGCTTAA 930	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
930	CTGTGCATGCTTAA 941	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
941	CTGTGCATGCTTAA 952	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
952	CTGTGCATGCTTAA 963	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
963	CTGTGCATGCTTAA 974	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
974	CTGTGCATGCTTAA 985	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
985	CTGTGCATGCTTAA 996	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
996	CTGTGCATGCTTAA 1007	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1007	CTGTGCATGCTTAA 1018	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1018	CTGTGCATGCTTAA 1029	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1029	CTGTGCATGCTTAA 1040	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1040	CTGTGCATGCTTAA 1051	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1051	CTGTGCATGCTTAA 1062	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1062	CTGTGCATGCTTAA 1073	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1073	CTGTGCATGCTTAA 1084	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1084	CTGTGCATGCTTAA 1095	BY727598	669 bp	mRNA	linear	EST 17-DEC-2002		
1095								

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

```

FEATURES             Location/Qualifiers
     source            1..669
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cdate="B73003H24"
                     /clone="B73003H24"
                     /issue_type="medulla oblongata"
                     /dev_stage="6 days neonate"
                     /lab_host="Dhl0B"
                     /clone_1db="RIKEN full-length enriched, 6 days neonate medulla oblongata"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGGAGAGAAGATCCAGACCTTTTTTTTTTTTNN 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGAATTCGGATTAAATTAATTAATCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN
Query Match      33.1%; Score 576.4; DB 4; Length 669;
Best Local Similarity 98.8%; Pred. No. 2,4e+161;
Matches 580; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY   1 ATGCCCTTTCATGTGGAAGACTGTGNGTACTTATTCCTCTTCACTCCTTATATTTCTG 60
DB   81 ATGCTTTCCACGTAGAAGACTGTGTGCTAATTAATCCTTTCACTCCTTATATTTCTG 140
QY   61 GTTGGAATATGCGGTGCATGGAAGAACCAAAAACGCGGCAACCAGAAAGCGCAGTAA 120
DB   141 GTTGGAATATGCGGTGCATGGAAGAACCAAAAACGCGGCAACCAGAAAGCGCAGTAA 200
QY   121 GCCATCATAGTCGGGGGCGGTGACATTGTTGTTGGTGGTGGTATTTTACCATGACCC 180
DB   201 GCCATCATAGTCGGGGGCGGTGACATTGTTGTTGGTGGTGGTATTTTACCATGACCC 260
QY   181 ACCGTGGTTGAGAGAGCTACATCAATGGGACGACAAGCAAGTATGGGCGCAGTTGT 240
DB   261 ACCGTGGTTGAGAGAGCTACATCAATGGGACGACAAGCAAGTATGGGCGCAGTTGT 320
QY   241 GGTTTACCTTGGGGCTCATGACCCATTTGATTTCTGTGAGTCTAATTTTAGTGCTG 300
DB   321 GGTTTACCTTGGGGCTCATGACCCATTTGATTTCTGTGAGTCTAATTTTAGTGCTG 380
QY   301 TTTTTCGGAACCTATAGCGTTCCAGAGGATATGTGCTATGTTAGCCCATTCAAACAG 360
DB   381 TTTTTCGGAACCTATAGCGTTCCAGAGGATATGTGCTATGTTAGCCCATTCAAACAG 440

```

```

QY 361 ATCTATGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTCT 420
Db 441 ATCTATGAAAGCGCATGGGTGGGCTGCTCTTCATCCCTGCACTGATGGAGAGATGTTCT 500
QY 421 TGGGCTGACGCAATTTTCTCTGCATTAGGGGGCCACCATTAAGCTGATCACTTATGATGGAT 480
Db 501 TGGGCTGACGCAATTTTCTCTGCATTAGGGGGCCACCATTAAGCTGATCACTTATGATGGAT 560
QY 481 GTGAACATATCGGTCATTTGCTCTGCACTCATTTGCATTTTATACCTTAAGTGGGTGGG 540
Db 561 GTGAACATATCGGTCATTTGCTCTGCACTCATTTGCATTTTATACCTTAAGTGGGTGGG 620
QY 541 CTCTACTCTGTGGCATATATCTGATGATGTTGTGCAGCTATTTCTGCATTTT 587
Db 621 CTCTACTCTGTGGCATATATCTGATGATGTTGTGCAGCTATTTCTGCATTTT 667

RESULT 9
LOCUS CK770440/c 882 bp mRNA linear EST 20-FEB-2004
DEFINITION 958613 MARC 1BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CK770440
VERSION CK770440.1 GI:42724534
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Peccora; Bovidae; Bovinae; Bos.
1 (bases 1 to 882)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Caessens,E., Wray,J.B., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
11282978
JOURNAL
PUBMED
COMMENT
Contact: Smith TP
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: 95 row: L column: 20
Seq primer: GTAATACGACCTCATATAGGG.
Location/Qualifiers
source
1..882
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
/note="Vector: PCMV SPORF6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
ORIGIN
Query Match 33.1%; Score 576.4; DB 5; Length 882;
Best Local Similarity 82.9%; Pred. No. 2.8e-161;
Matches 658; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

950 ACATATATCTCCGATCGTTTGCAGTACTTGCCTGTGATCACTCTTCTTTGGGC 1009
|||||
882 ACATATATCTCCGATCGTTTGCAGTACTTGCCTGTGATCACTCTTCTTTGGGC 823
|||||
1010 TTGGTGTGTTTCACTGTGTCATGTCCTCTCAGCTGACCGTCATCCCTGTGGGAGATT 1069
|||||

```

```

Db 822 TTGGAGCCGTTTCTGCTCTGCTGATGTCCTCATGAGCATTTCTTCCATCTTGCAGAAATT 763
QY 1070 CTATGTTGCTCGGAATATCTACAGCTTTCTCTTCCAGCAAAATSCATCAGACAGAAA 1129
|||||
762 CGATGTTGCTCGGAATATCTACAGCTTTCTCTTCCAGCAAAATSCATCAGACAGAAA 703
|||||
1130 TTGTGTGTCATGAGATCACTGCTGTTGTGTCGAGCATCTGCAACAGCATGCTT 1189
|||||
702 TAGTCTGGGTATGGGATCAAGGATTTGTTGTTGGAGCTTCTGCGATGACCAATGGGCT 643
|||||
1130 TGCTGACGAAGATGATGAGGCTCTGTGTAAGCTGAGCTGAGCTTGTATCATCATCA 1249
|||||
642 TGTACCAAGACGATGATGAGGCTCTGTGTAAGCTGAGCTGAGCTTGTATCATCATCA 583
|||||
1250 TCTTCCACAGCTGCTGTGTAAGCTTCACTCAAAAGAACCAACTATGAGGAGTTG 1309
|||||
582 TCTTCCACAGCTGCTGTGTAAGCTTCACTCAAAAGAACCAACTATGAGGAGTTG 523
|||||
1310 CTGTTATATTTTGGACTATTTCTGAGAAATTAAGTGAAGAGCATATCTATCTTGC 1369
|||||
522 CAGTTACATCTCGGGCTTTTCTGCGAGTCACTGCGGGAGCCCTAACCTGAACCTGC 463
|||||
1370 AGCCTTATCTTCTCAACCTGTTATTAATCTGACAAAGATGATATATCAATGAGGT 1429
|||||
462 AGCCTTATCTTCTCAACCTGTTATTAATCTGACAAAGATGATATATCAATGAGGT 403
|||||
1430 TCCCATTTAAACCTCTCCATATGTTACCTCATTTCTTACCAACATTTGTGTTCTTATC 1489
|||||
402 TCCCATTTAAACCTCTCCATATGTTACCTCATTTCTTACCAACATTTGTGTTCTTATC 343
|||||
1490 TAGCCAGTATCTATTTGAAAGTGAACCTTGCCTCCAAATTTAGATGATTTGATGCTG 1549
|||||
342 TAGCCAGTATCTATTTGAAAGTGAACCTTGCCTCCAAATTTAGATGATTTGATGCTG 283
|||||
1550 TTGTGCAAGGACAGTGAAGAACATGAGCAAGACCATTTAGTCAAGAAATGAATA 1609
|||||
282 TTGTGCAAGGACAGTGAAGAACATGAGCAAGACCATTTAGTCAAGAAATGAATA 223
|||||
1610 TCAAAATTAATGAACTTGCACCTGTGAACCTCGCAGAGCTTAACCTCACTCACTT 1669
|||||
222 TTAACCTGATGAACTTGCACCTGTGAACCTCGCAGAGCTTAACCTCACTCACTT 163
|||||
1670 TCACCAATPAGAGAGCCCTCTTGAATGTTGATTCAGATCCGAGAGGGTCTGGACTGAAG 1729
|||||
162 TCACCAATPAGAGAGCCCTCTTGAATGTTGATTCAGATCCGAGAGGGTCTGGACTGAAG 103
|||||
1730 ATATTTTCAATGA 1743
|||||
102 ATATTTTCAATGA 89
|||||

RESULT 10
LOCUS CT403380 828 bp DNA linear GSS 03-NOV-2005
DEFINITION Sus scrofa genomic clone CH242-157K9, genomic survey sequence.
ACCESSION CT403380
VERSION CT403380.1 GI:80060743
KEYWORDS GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.
1 (bases 1 to 828)
Humphray,S.J., Plumb,R.W. and Durham,J.L.
Direct Submission
Submitted (01-NOV-2005) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humphray@sanger.ac.uk Unpublished
This sequence was generated from the SPE end of BAC 157K9. 157K9 is
part of the CHORI-242 BAC Library created by P. de Jong. Further
details: http://www.sanger.ac.uk/Projects/S_sus/

```



ORIGIN BamHI"

Query Match 23.5%; Score 409.6; DB 7; Length 650;  
Best Local Similarity 96.7%; Pred. No. 4.2e-111;  
Matches 437; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1 ATGCTTTCATGTCGAAAGAGCTGATGCTATATCTCTTCTTACCTCTTATATTTCTG 60  
DB 201 ATGCTTTCACGTCGAAAGAGCTGATGCTATATCTCTTCTTACCTCTTATATTTCTG 260  
QY 61 GTTGAATATGCGGCGCATGGAACCAAAACACGCGAACCAGGAAGGCGCATGA 120  
DB 261 GTTGAATATGCGGCGCATGGAACCAAAACACGCGAACCAGGAAGGCGCATGA 320  
QY 121 GCCATCATATGTCGCGGCGCATGATGCTGTTGTTGTTGTTGTTTACATGACAGCC 180  
DB 321 GCCATCATATGTCGCGGCGCATGATGCTGTTGTTGTTGTTTACATGACAGCC 380  
QY 181 ACCTGGTTGAGAGAGCTACATCAATGGACAGAGAGAGTATGGCCAGGTTGT 240  
DB 381 ACCTGGTTGAGAGAGCTACATCAATGGACAGAGAGAGTATGGCCAGGTTGT 440  
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATCTCTGAGCTATATTTAGGTCG 300  
DB 441 GGTCTAGCTTGGGCTCATGCAACCATTTGATATCTCTGAGCTATATTTAGGTCG 500  
QY 301 TTTTGGGAAACCTATGCGGTCGCAAGGATATGACTATGTAGACCATTCACAG 360  
DB 501 TTTTGGGAAACCTATGCGGTCGCAAGGATATGACTATGTAGACCATTCACAG 559  
QY 361 ATCTATGAAAGCGCATGGGTCGCTCTTCATCCTGCACTGATGGAGAGATGTC 420  
DB 560 ATCTATGAAAGCGCATGGGTCGCTCTTCATCCTGCACTGATGGAGAGATGTC 619  
QY 421 TGGGCTGCAAAATTTCTCTGCTATTAAGGGC 452  
DB 620 TGGGCTGCAAAATTTCTCTGCTATTAAGGGC 650

RESULT 12  
AG157499 672 bp DNA linear GSS 09-JAN-2002  
LOCUS Pan troglodytes DNA, clone: RP43-022H02.T7, genomic survey  
DEFINITION  
ACCESSION AG157499.1 GI:16687177  
VERSION AG157499.1  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM  
REFERENCE  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of Library RP43-022H02  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 672)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/),  
Tel: 81-45-503-9111, Fax: 81-45-503-9170  
COMMENT Clones are derived from the chimpanzee BAC library RP43-022H02  
of clone tracking errors.  
PRIMERS  
Sequencing: T7  
LIBRARY

FEATURES  
source

ORIGIN

Query Match 23.3%; Score 406.8; DB 14; Length 672;  
Best Local Similarity 84.9%; Pred. No. 3e-110;  
Matches 479; Conservative 0; Mismatches 82; Indels 3; Gaps 2;

QY 1110 AATGCAATGACAAAGAAATGTTGGTCAATGAGATCACTGCTGTTGTCGAGC 1169  
DB 111 ACAGGCTTCGACAAAGAAATGTTGGTCAATGAGATCACTGCTGTTGTCGAGC 170  
QY 1170 ATCTGCAAGGCAATGCTTGTGAGAGAGCTGTATGAGGCTGTGATCCTGAGCTC 1229  
DB 171 ATCTGCAAGGCAATGCTTGTGAGAGAGCTGTATGAGGCTGTGATCCTGAGCTC 230  
QY 1230 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGATCTTCAATCAAGAAC 1289  
DB 231 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGATCTTCAATCAAGAAC 290  
QY 1290 CAACCTATGAGGAGGTCGTTGATATTTTGGACTATCTTCCGAGATTAACGAGG 1349  
DB 291 CAACCTATGAGGAGGTCGTTGATATTTTGGACTATCTTCCGAGATTAACGAGG 350  
QY 1350 AGAGCATATCTATCTTCAAGGCTTCACTTCACTGATCTTCACTGATCTTCACTG 1409  
DB 351 AGAGCATATCTATCTTCAAGGCTTCACTTCACTGATCTTCACTGATCTTCACTG 410  
QY 1410 TGGTATATATATATGAAATTTCCATTTAAACCTCTCTCATGTTACTCTTAC 1469  
DB 411 TGGTATATATATATGAAATTTCCATTTAAACCTCTCTCATGTTACTCTTAC 470  
QY 1470 CAACCTTGTGTTTCTTATCTTCAAGGCTTCACTTCACTGATCTTCACTGAT 1528  
DB 471 CAACCTTGTGTTTCTTATCTTCAAGGCTTCACTTCACTGATCTTCACTGAT 530  
QY 1529 AATTGATGATATTTGATGCTGTTGCGCAAGGACAGTGAAGAGACATGACAGCA 1588  
DB 531 AATTGATGATATTTGATGCTGTTGCGCAAGGACAGTGAAGAGACATGACAGCA 590  
QY 1589 TTCTAGTCAGAAATGAAATATCAATTAATGAATTTGACCTGTGAAACCTGGCAGA 1648  
DB 591 TTTTGTCAAAAATGAAA--TATAATTAATGATGACCTTGTGAAGCAGCAGACGA 648  
QY 1649 GCCTAACCTCAGTTCACTTTCA 1672  
DB 649 ACATGACCTCAGCTTAACCTTTCA 672

RESULT 13  
DA361315 548 bp mRNA linear EST 05-NOV-2005  
LOCUS DA361315 BRSTN2 Homo sapiens cDNA clone BRSTN2007643 5', mRNA  
DEFINITION  
ACCESSION DA361315  
VERSION DA361315.1 GI:80797290  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
TITLE  
COMMENT 1 (bases 1 to 548)



```

Db      | 205 | GGCATCATAGTGGTGGCCGAGATATGTTATGTTGGTGGATTTACCATGACAGCT 264
Oy      | 181 | ACCTGGTGGAGAGGCTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db      | 265 | ACCTGGTGGAGAGAGGATATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
Oy      | 241 | GATCTAGCTGGGCTCATGACACCATTTGAGATTTCTGAGTCTAAATTTAGTGGTCTG 300
Db      | 325 | GGCTAGCTGGGCTCAGGACCAATTTGATTTCTTTAGTCTATTTAGTGGCTG 384
Oy      | 301 | TTTTGGGAAACCATGATGCTTCCAGGATATGAGATGATGATGATGATGATGATGAT 360
Db      | 385 | TTTTGGGAAACCATGATGCTTCCAGGATATGAGATGATGATGATGATGATGATGAT 444
Oy      | 361 | ATCTATGAAAGCGATGGTGGGCTGCTCTTCATCCCTGACATGATGAGAGAGATGTTT 420
Db      | 445 | ATCTATGAAAGCGATGGGCTGCTCTTCATCCCTGACATGATGAGAGAGATGTTT 504
Oy      | 421 | TGGGCTGACAGCAATTTTCTCTGATTTAGGGGCGACCATGAGCGTATCATTTGATGAT 480
Db      | 505 | TGGGCTGACAGCAATTTTCTCTGATTTAGGGGCGACCATGAGCGTATCATTTGATGAT 564
Oy      | 481 | GTGACATATTC 491
Db      | 565 | ATGACATTTTC 575

```

## RESULT 15

```

BE233479      516 bp      mRNA      linear      EST 10-JUL-2000
LOCUS      BE233479      139685      MABC      1P1G      Sus      scrofa      cDNA 5', mRNA sequence.
DEFINITION      BE233479.1      GI:9018197
VERSION      BE233479.1      GI:9018197
KEYWORDS      EST.
SOURCE      Sus scrofa (pig)
ORGANISM      Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

```

```

REFERENCE      1 (bases 1 to 516)
AUTHORS      Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
              Vallet, J., Wise, T., Rohrer, G.A., Ferreira, G., Sultana, R.,
              Queckenbush, J., and Keeler, J.W.
              Porcine gene discovery by normalized cDNA-library sequencing and
              EST cluster assembly
              Mamm. Genome 13 (8), 475-478 (2002)
JOURNAL      12226715
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366
              Fax: 402 762 4390
              Email: smith@mail.marc.usda.gov
              Single pass sequencing. Bases called and alt trimmed with phred
              v0.980904.e. Vector identified by cross_match with the -mscore 18
              and -mismatch 12 options.
              PCR Primers
              FORWARD: AGGAAACAGCTATGACCAT
              BACKWARD: GTTTCCAGTCAGCAGC
              Plate: 75 row: G column: 12
              Seq primer: ATTAGTGACATATAG.

```

```

FEATURES
SOURCE
1..516
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="Vector: PCMW SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

```

## ORIGIN

```

Query Match      22.4%; Score 389.6; DB 7; Length 516;
Best Local Similarity 84.7%; Pred. No. 4e-105;
Matches 437; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```

```

Oy      | 1067 | GTTCTATGTTTGGCTCGCAATATCTACAGCTTTCTCTTACGCAAAATGATCATGACAG 1126
Db      | 1      | GTTCTATGTTTGGCTCGCAATATCTACAGCTTTCTCTTACGCAAAATGATCATGACAG 60
Oy      | 1127 | AAATGTGTGGGTCATGAGATCACTGCTGTTGGTGGAGCATCTGCAACAGCATG 1186
Db      | 61      | AGATGCTGTGGGTCATGAGATCACTGCTGTTGGTGGAGCATCTGCAACAGCATG 120
Oy      | 1187 | CTTTGTGAGCAAGACTGTGTATGAGGCTCTGTTGACTGTGAGCTTGTGACCTTGTATCA 1246
Db      | 121      | CTTTGTGAGCAAGACTGTGTATGAGGCTCTGTTGACTGTGAGCTTGTGACCTTGTATCA 180
Oy      | 1247 | TCATCTTCCACAGCTGTCTGTGTACTCTTCATCAAGAAACCAACCTATGAGGAG 1306
Db      | 181      | TTATCTTCCACAGCTGTCTGTGTACTCTTCATCAAGAAACCAACCTATGAGGAG 240
Oy      | 1307 | TTGCTGTTATATTTTGGACATTTCTGAGATTTACTGAGAGAGAGCATATCTACT 1366
Db      | 241      | TGGCAGGATCATTTGCTGCTGCTTTCTGAGGATACCGGTGAGAGCATATCTACT 300
Oy      | 1367 | TGCAGCCCTTAATCTTCTACCCCTGTTATTACTGACAAAGATGATATATACATCA 1426
Db      | 301      | TGCAGCCCTTAATCTTCTACCCCTGTTATTACTGACAAAGATGATATATATACATCA 360
Oy      | 1427 | GGTTCCTCAATTTAAACTCTCTTCATGTTACTTATTTTACCAACATTTGTGTTCTT 1486
Db      | 361      | GATTCCTCAATTTAAACTCTCTTCATGTTACTTATTTTACCAACATTTGTGTTCTT 420
Oy      | 1487 | ATCTAGCAAGATATCTTAAAGTGAACCTTGCCCAAAATTTAGATGATTTGATG 1546
Db      | 421      | ATCTAGCAAGATATCTTAAAGTGAACCTTGCCCAAAATTTAGATGATTTGATG 480
Oy      | 1547 | CTGTTGTGCAAGGACAGTGAAGAGACATGAGCA 1582
Db      | 481      | CTGTTGTGCAAGGACAGTGAAGAGACATGAGATA 516

```

```

Search completed: July 10, 2006, 16:31:14
Job time : 8719 secs

```

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:18 ; Search time 1996 Seconds  
(without alignments)  
10730.134 Million cell updates/sec

Title: US-10-724-806-3

Sequence: 1 atgccttcacatggaag.....ctgaataatcacaatga 1743

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 3784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1743	100.0	1743	3	US-09-911-077A-3
2	1743	100.0	1743	3	US-09-911-077A-23
3	1743	100.0	1743	3	US-10-724-806-3
4	1743	100.0	1743	9	US-10-724-806-23
5	1616.6	92.7	4904	9	US-09-911-077A-5
6	1616.6	92.7	4904	9	US-10-724-806-5
7	1375	78.9	1743	9	US-09-911-077A-1
8	1375	78.9	1743	9	US-10-724-806-1
9	1375	78.9	1813	9	US-09-911-077A-9
10	1375	78.9	1813	9	US-10-724-806-9
11	502.8	28.8	26323	15	US-11-005-030-1
12	502.8	28.8	119040	9	US-09-911-077A-19
13	502.8	28.8	119040	9	US-10-724-806-19
14	502.8	28.8	142299	9	US-09-911-077A-14
15	502.8	28.8	142299	9	US-10-724-806-14
16	388.8	22.3	1833	8	US-10-241-784-1
17	355.8	20.4	1985	3	US-09-911-077A-7

18	355.8	20.4	1985	9	US-10-724-806-7	Sequence 7, Appl1
19	278.8	16.0	1729	13	US-11-097-143-41594	Sequence 41594, A
20	237.6	13.6	1461	3	US-09-974-300-501	Sequence 501, App
21	178	10.2	4223	13	US-11-097-143-41593	Sequence 41593, A
22	145.6	8.4	1474	4	US-09-925-0654-717810	Sequence 717810, A
23	145.6	8.4	1474	5	US-09-925-0654-717810	Sequence 717810, A
24	145.6	8.4	119040	9	US-09-911-077A-19	Sequence 19, Appl1
25	145.6	8.4	119040	9	US-10-724-806-19	Sequence 19, Appl1
26	145.6	8.4	142299	9	US-09-911-077A-14	Sequence 14, Appl1
27	145.6	8.4	142299	9	US-10-724-806-14	Sequence 14, Appl1
28	132.4	7.6	455	3	US-09-864-761-1898	Sequence 1898, App
29	132.4	7.6	1094	10	US-10-450-763-11997	Sequence 11997, A
30	95	5.5	581	4	US-09-925-0654-760066	Sequence 760066, A
31	95	5.5	581	5	US-09-925-0654-760066	Sequence 760066, A
32	94.6	5.4	943	6	US-10-027-632-120553	Sequence 120553, A
33	94.6	5.4	943	7	US-10-027-632-120553	Sequence 120553, A
34	93.8	5.4	600	10	US-10-972-079-24268	Sequence 24268, A
35	65	3.7	65	3	US-09-908-975-26842	Sequence 26842, A
36	55.6	3.2	70	10	US-10-957-432-58	Sequence 58, App1
37	53.6	3.1	60	3	US-09-908-975-10249	Sequence 10249, A
38	52.8	3.0	96	3	US-09-864-761-18589	Sequence 18589, A
39	41.4	2.4	969	3	US-09-925-299-182	Sequence 182, App
40	41.4	2.4	969	3	US-09-925-299-182	Sequence 182, App
41	41.2	2.4	351	10	US-10-779-543-11719	Sequence 11719, A
42	39.8	2.3	3091	10	US-10-764-420-1219	Sequence 1219, App
43	39.8	2.3	3100	16	US-11-136-527-592	Sequence 592, App
44	39.4	2.3	933	8	US-10-335-977-516	Sequence 516, App
45	39.4	2.3	1338	8	US-10-335-977-517	Sequence 517, App

#### ALIGNMENTS

RESULT 1  
US-09-911-077A-3  
Sequence 3, Application US/0911077A  
Publication No. US20030114399A1  
GENERAL INFORMATION:  
APPLICANT: BLAKELY, RANDY D.  
APPLICANT: APPARISUNDARAM, SUBRAMANIAM  
APPLICANT: FERGUSON, SHAWN  
TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
FILE REFERENCE: VBLT:008US  
CURRENT APPLICATION NUMBER: US/09/911, 077A  
CURRENT FILING DATE: 2001-07-23  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 1743  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1743)  
US-09-911-077A-3

Query Match 100.0% Score 1743; DB 3; Length 1743;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTTTCATGAGAGAGCTGATGCTATATCTTCTTACCTCTTATATTTCTG	60
DB	1	ATGCTTTCATGAGAGAGCTGATGCTATATCTTCTTACCTCTTATATTTCTG	60
QY	61	GTGGAAATATGGCTGATGAGAAACCAAAACAGGCGAACCCAGAGGCGAGTGA	120
DB	61	GTGGAAATATGGCTGATGAGAAACCAAAACAGGCGAACCCAGAGGCGAGTGA	120
QY	121	GCATCATATGCGGGGCGCGTGAATGTTGTTGTTGTTGTTTACATGACAGCC	180
DB	121	GCATCATATGCGGGGCGCGTGAATGTTGTTGTTGTTTACATGACAGCC	180
QY	181	ACCTGGTTGAGAGAGGCTTACATCAATGGAGACGACGAGACGAGTGTGT	240

```
Db      ||||| 181 ACCTGGTGTGGAGGCTACATCATGGAGCAGAGCAGTGTATGGCCAGGTGT 240
Oy      ||||| 241 GGTACGCTGGGCTCATGCCACCATGGATATTCCTGAGCTATTTTAGTGTCTG 300
Db      ||||| 241 GGTACGCTGGGCTCATGCCACCATGGATATTCCTGAGCTATTTTAGTGTCTG 300
Oy      ||||| 301 TTTTTCGGAACCTATGGCTTCCAAAGGATATGTACTATGTAGACCATTCACAG 360
Db      ||||| 301 TTTTTCGGAACCTATGGCTTCCAAAGGATATGTACTATGTAGACCATTCACAG 360
Oy      ||||| 361 ATCTATGAAAAGCGATGGGTGGCTGCTTCATCCCTGCGACTGATGGGAGAGATGTC 420
Db      ||||| 361 ATCTATGAAAAGCGATGGGTGGCTGCTTCATCCCTGCGACTGATGGGAGAGATGTC 420
Oy      ||||| 421 TGGGCTGAGCAATTTCTCTGCTATTAAGGGGCCACATCAGCGTATCATGTATGAT 480
Db      ||||| 421 TGGGCTGAGCAATTTCTCTGCTATTAAGGGGCCACATCAGCGTATCATGTATGAT 480
Oy      ||||| 481 GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTTATACCTTATGGGTGG 540
Db      ||||| 481 GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTTATACCTTATGGGTGG 540
Oy      ||||| 541 CTCTACTCTGTCGCAATATCTGATGTGTCCAGCTATTTCTGATTTTATAGACTGTGG 600
Db      ||||| 541 CTCTACTCTGTCGCAATATCTGATGTGTCCAGCTATTTCTGATTTTATAGACTGTGG 600
Oy      ||||| 601 ATCAGTGTCCCTTTGGCCCTGTCACATCCTGCACTCACCGCATGGATTCACAGCTGTG 660
Db      ||||| 601 ATCAGTGTCCCTTTGGCCCTGTCACATCCTGCACTCACCGCATGGATTCACAGCTGTG 660
Oy      ||||| 661 CATGCTAAATACCAAGTCTCTGGCTGGGACCATTTGAATCACTTGAAGTCTTACCTGG 720
Db      ||||| 661 CATGCTAAATACCAAGTCTCTGGCTGGGACCATTTGAATCACTTGAAGTCTTACCTGG 720
Oy      ||||| 721 CTGTATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCCAGAG 780
Db      ||||| 721 CTGTATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCCTACTTCCAGAG 780
Oy      ||||| 781 GTCTCTCTTCATCCTTCAGCAGCACTATGCTCAGATATGCTCTCTCTGCACTTTTGGG 840
Db      ||||| 781 GTCTCTCTTCATCCTTCAGCAGCACTATGCTCAGATATGCTCTCTCTGCACTTTTGGG 840
Oy      ||||| 841 TGCCCTGTGATGCTCTACCCGCCCATATGCAATGAGAGCTATTTGAGCTTCCACAGCTGG 900
Db      ||||| 841 TGCCCTGTGATGCTCTACCCGCCCATATGCAATGAGAGCTATTTGAGCTTCCACAGCTGG 900
Oy      ||||| 901 AACGAGACTGCTAGGGGATCCAGATCCCAAGACTAAGAGAGAGAGACATGATTC 960
Db      ||||| 901 AACGAGACTGCTAGGGGATCCAGATCCCAAGACTAAGAGAGAGAGACATGATTC 960
Oy      ||||| 961 CGGATCGTTCAGATCTCTGCTGCTGTATCATCTCTTCTTGGGCTGTGTGTGT 1020
Db      ||||| 961 CGGATCGTTCAGATCTCTGCTGCTGTATCATCTCTTCTTGGGCTGTGTGTGT 1020
Oy      ||||| 1021 TCAGCTGCTGTATGCTCTCAGCTGATCTGTCTATCTGTGGCGAGTTCATATGTTTGT 1080
Db      ||||| 1021 TCAGCTGCTGTATGCTCTCAGCTGATCTGTCTATCTGTGGCGAGTTCATATGTTTGT 1080
Oy      ||||| 1081 CGGATATCTACACAGCTTCTTCCAGCAAAATGCAATGCAAGAGAAATTTGTGGGCT 1140
Db      ||||| 1081 CGGATATCTACACAGCTTCTTCCAGCAAAATGCAATGCAAGAGAAATTTGTGGGCT 1140
Oy      ||||| 1141 ATGAGGATCATGCTGCTGTGTGGAGCATGTGCAACGCGCATGGCTTGTGTCAGAG 1200
Db      ||||| 1141 ATGAGGATCATGCTGCTGTGTGGAGCATGTGCAACGCGCATGGCTTGTGTCAGAG 1200
Oy      ||||| 1201 ACTGTGTATGGGCTCTGTGACCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260
Db      ||||| 1201 ACTGTGTATGGGCTCTGTGACCTGAGCTCTGACCTTGTCTACATCATCTTCCACAG 1260
Oy      ||||| 1261 CTGCTGTGTATCTCTTCAACAGAACCAACATATAGGGGCGAGTTCGTTATAT 1320
Db      ||||| 1261 CTGCTGTGTATCTCTTCAACAGAACCAACATATAGGGGCGAGTTCGTTATAT 1320
```

```
Db      ||||| 1261 CTGCTGTGTATCTCTTCAACAGAACCAACATATAGGGGCGAGTTCGTTATAT 1320
Oy      ||||| 1321 TTTGACATATCTCTGAGAAATTAATGAGAGAGCCATATCTATCTTSCAGCCCTTAATC 1380
Db      ||||| 1321 TTTGACATATCTCTGAGAAATTAATGAGAGAGCCATATCTATCTTSCAGCCCTTAATC 1380
Oy      ||||| 1381 TTCTACCTGTATTAATCTGACAGAAATGATATATCAATCAGAGTTCCCATTTAA 1440
Db      ||||| 1381 TTCTACCTGTATTAATCTGACAGAAATGATATATCAATCAGAGTTCCCATTTAA 1440
Oy      ||||| 1441 ACTCTCTCAATGTTACCTCATCTTCAACAAATTTGTGTCTTATCTAGCCCAAGAT 1500
Db      ||||| 1441 ACTCTCTCAATGTTACCTCATCTTCAACAAATTTGTGTCTTATCTAGCCCAAGAT 1500
Oy      ||||| 1501 CTATTTGAAATGGAACCTTGCCTTCAAAATTAATGATGATTTGATGCTGTGTGCAAG 1560
Db      ||||| 1501 CTATTTGAAATGGAACCTTGCCTTCAAAATTAATGATGATTTGATGCTGTGTGCAAG 1560
Oy      ||||| 1561 CACAGTGAAGAAACATGAGCAAGACATTTCTAGTCAAGAAATGAAATATCAATTAAT 1620
Db      ||||| 1561 CACAGTGAAGAAACATGAGCAAGACATTTCTAGTCAAGAAATGAAATATCAATTAAT 1620
Oy      ||||| 1621 GAACCTTGACCTGTGAACCTGCGAGGCTTAACCTCAGTTCAACTTCAACCAATAG 1680
Db      ||||| 1621 GAACCTTGACCTGTGAACCTGCGAGGCTTAACCTCAGTTCAACTTCAACCAATAG 1680
Oy      ||||| 1681 GAGGCCCTCTGTATGTTATTCAGTCCGAGGGGTCTGGGACTGAAGATTAATTTCA 1740
Db      ||||| 1681 GAGGCCCTCTGTATGTTATTCAGTCCGAGGGGTCTGGGACTGAAGATTAATTTCA 1740
Oy      ||||| 1741 TGA 1743
Db      ||||| 1741 TGA 1743

RESULT 2
US-09-911-077A-23
; Sequence 23, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBLT:0080US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-09-911-077A-23

Query Match      100.0%; Score 1743; DB 3; Length 1743;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      ||||| 1 ATGCTTTTCATGTGGAAGACTGTGTACTATATATCTCTTCTTCACTTATATTTCTG 60
Db      ||||| 1 ATGCTTTTCATGTGGAAGACTGTGTACTATATATCTCTTCTTCACTTATATTTCTG 60
Oy      ||||| 61 GTTGAATATATGGGCTGCAATGGAACCAAAACAGCGGCAACCCAGAGAGCGCAGTAA 120
Db      ||||| 61 GTTGAATATATGGGCTGCAATGGAACCAAAACAGCGGCAACCCAGAGAGCGCAGTAA 120
Oy      ||||| 121 GCCATCATATGTCGGGGGCGTGACATGTTGTGTGTGTGTGTGTATACCATGACAGCC 180
Db      ||||| 121 GCCATCATATGTCGGGGGCGTGACATGTTGTGTGTGTGTGTGTATACCATGACAGCC 180
```



```
QY 181 ACCTGGGTGAGAGAGCTACATCAATGGGAGACGAGATGATGGGCCAGGTTGT 240
DB 181 ACCTGGGTGAGAGAGCTACATCAATGGGAGACGAGATGATGGGCCAGGTTGT 240
QY 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCAATTTTGGTGTCTG 300
DB 241 GGTCTAGCTTGGGCTCATGCAACCATTTGATATTTCTGAGTCAATTTTGGTGTCTG 300
QY 301 TTTTTCGCAAACTATGCGCTTCAAGGATATGATCTATGTTAACCATTCAACAG 360
DB 301 TTTTTCGCAAACTATGCGCTTCAAGGATATGATCTATGTTAACCATTCAACAG 360
QY 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCACTGCACTGATGAGAGAGATGTC 420
DB 361 ATCTATGAAAGGAGATGGGTGGGCTCTTCACTGCACTGATGAGAGAGATGTC 420
QY 421 TGGGCTGAGCAATTTTCTGCAATTTAGGGCCACCATGAGCGTGAATCAATGATGAT 480
DB 421 TGGGCTGAGCAATTTTCTGCAATTTAGGGCCACCATGAGCGTGAATCAATGATGAT 480
QY 481 GTGACATATCGGTCAATGCTCTGCACTATTCGCAATTTCTTATACCTTATGGGTGG 540
DB 481 GTGACATATCGGTCAATGCTCTGCACTATTCGCAATTTCTTATACCTTATGGGTGG 540
QY 541 CTCTACTCTGTGGCATATATCTGATGTTGTCAGACTATTCGCAATTTTATAGGACTGTG 600
DB 541 CTCTACTCTGTGGCATATATCTGATGTTGTCAGACTATTCGCAATTTTATAGGACTGTG 600
QY 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCACTGACGACATCGGATTCACAGCTGTG 660
DB 601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGCACTGACGACATCGGATTCACAGCTGTG 660
QY 661 CATGCTAAATACCAAGAGTCCCTGGCTGGGAGAACCATTTGATCACTGATCACTGTG 720
DB 661 CATGCTAAATACCAAGAGTCCCTGGCTGGGAGAACCATTTGATCACTGATCACTGTG 720
QY 721 CTGTGATATTTTCTGATTTGATGCTGGGTGGAATCCCATGGACGCACTACTTCCAGAG 780
DB 721 CTGTGATATTTTCTGATTTGATGCTGGGTGGAATCCCATGGACGCACTACTTCCAGAG 780
QY 781 GTCTCTCTTCACTCTCAGCACTATGCTCAGAGTACTGTCTTCTGCGAGCTTTTGGG 840
DB 781 GTCTCTCTTCACTCTCAGCACTATGCTCAGAGTACTGTCTTCTGCGAGCTTTTGGG 840
QY 841 TGCCTGTGATGCTCTACCCGCAATATGATGAGCTATTTGAGCTTCCACAGACTG 900
DB 841 TGCCTGTGATGCTCTACCCGCAATATGATGAGCTATTTGAGCTTCCACAGACTG 900
QY 901 AACCAAGCTGCTAGGGGATCCAGATCCCAAGATTAAGGAGAGACATGATTTCTC 960
DB 901 AACCAAGCTGCTAGGGGATCCAGATCCCAAGATTAAGGAGAGACATGATTTCTC 960
QY 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTTCTTCTTGGGCTTGTGCTGT 1020
DB 961 CCGATGCTTCTGAGTACCTGCTGCTGCTGCTGCTGCTTCTTCTTGGGCTTGTGCTGT 1020
QY 1021 TCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CCGAATATCTACAGGCTTCTTCAAGCAAAATGATCAGACAAAGAAATGTGTGGTCT 1140
DB 1081 CCGAATATCTACAGGCTTCTTCAAGCAAAATGATCAGACAAAGAAATGTGTGGTCT 1140
QY 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 ATGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
```

```
QY 1261 CTGCTCTGTGTACTCTTCAATCAAGAACCAACCTTATGGGCACTTCTGCTTATTT 1320
DB 1261 CTGCTCTGTGTACTCTTCAATCAAGAACCAACCTTATGGGCACTTCTGCTTATTT 1320
QY 1321 TTTGACATATCTCTGAGAAATTTCTGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1380
DB 1321 TTTGACATATCTCTGAGAAATTTCTGAGAGAGCCATATCTATCTTACAGCCCTTAATC 1380
QY 1381 TTCTACCTGCTTATTTACTCTGACAAAGATGATATCAATCAGAGGTTCCATTAA 1440
DB 1381 TTCTACCTGCTTATTTACTCTGACAAAGATGATATCAATCAGAGGTTCCATTAA 1440
QY 1441 ACTCTCTCCATGTTATCTCTTATTTACCAATTTGCTTCTTATCTAGCAAGTAT 1500
DB 1441 ACTCTCTCCATGTTATCTCTTATTTACCAATTTGCTTCTTATCTAGCAAGTAT 1500
QY 1501 CTATTTGAAAGTGAAACCTTGGCTCCAAATTTAGATGATTTGATCTGTTGCGAAG 1560
DB 1501 CTATTTGAAAGTGAAACCTTGGCTCCAAATTTAGATGATTTGATCTGTTGCGAAG 1560
QY 1561 CACAGTGAAGAGAACTGGAACAAGACATTTAGTCAAGAAATGAAATCAATTAAT 1620
DB 1561 CACAGTGAAGAGAACTGGAACAAGACATTTAGTCAAGAAATGAAATCAATTAAT 1620
QY 1621 GAACCTTGCACTGTGAAACCTGCGCAGAGCTTAACCTCAGTTCAACCTTCAACAA 1680
DB 1621 GAACCTTGCACTGTGAAACCTGCGCAGAGCTTAACCTCAGTTCAACCTTCAACAA 1680
QY 1681 GAGGCTCTCTGATGTTATTTCACTGCGAGGGGTCTGGGATGAAATTTTACAA 1740
DB 1681 GAGGCTCTCTGATGTTATTTCACTGCGAGGGGTCTGGGATGAAATTTTACAA 1740
QY 1741 TGA 1743
DB 1741 TGA 1743
```

## RESULT 3

```
US-10-724-806-3
; Sequence 3, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARSDUNDAM, SUBRAMANIAM
; APPLICANT: FERGIUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1743)
US-10-724-806-3
```

```
Query Match 100.0%; Score 1743; DB 9; Length 1743;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTTTCATGATGAGAGAGCTGATGATATCTCTTCACTCTTATATTTCTG 60
DB 1 ATGCTTTTCATGATGAGAGAGCTGATGATATCTCTTCACTCTTATATTTCTG 60
QY 61 GTTGAATATGGGCTGCTGAGAAACCAAAACAGGGCAACCCAGAGGCGCAGTGA 120
DB 61 GTTGAATATGGGCTGCTGAGAAACCAAAACAGGGCAACCCAGAGGCGCAGTGA 120
```

121 GCCATCATAGTCGGGGGCGGTGACATGGTGTGTTGTTGTTGTTTACCATGACAGC 180  
121 GCCATCATAGTCGGGGGCGGTGACATGGTGTGTTGTTGTTGTTTACCATGACAGC 180  
181 ACCTGGGTTGAGAGGCTACATCAATGGGACAGAGAGAGAGAGAGAGAGAGAGAG 240  
181 ACCTGGGTTGAGAGGCTACATCAATGGGACAGAGAGAGAGAGAGAGAGAGAGAG 240  
241 GGTACGCTGGGCTCATGCAACCATGGATATCTCTGAGCTTAATTTTAAAGGTCTG 300  
241 GGTACGCTGGGCTCATGCAACCATGGATATCTCTGAGCTTAATTTTAAAGGTCTG 300  
301 TTTTTCGGAAAACCTATGGGTTCCAAAGGATATGATCTATGTTAGACCATTAACAG 360  
301 TTTTTCGGAAAACCTATGGGTTCCAAAGGATATGATCTATGTTAGACCATTAACAG 360  
361 ATCTATGAGAAAGCGATGGGTGGGCTGCTCTTCAATCCCTGCACTGATGGGAGAGAT 420  
361 ATCTATGAGAAAGCGATGGGTGGGCTGCTCTTCAATCCCTGCACTGATGGGAGAGAT 420  
421 TGGGCTGAGCAATTTTCTGCTCATTTAGGGGCGCACCATGAGGCTATCATGATGAT 480  
421 TGGGCTGAGCAATTTTCTGCTCATTTAGGGGCGCACCATGAGGCTATCATGATGAT 480  
481 GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTAAACCTTAAAGTGG 540  
481 GTGAACATATCGGCTATGCTCTGCACTCATTTGCCATTTCTTAAACCTTAAAGTGG 540  
541 CTCTCTCTGAGGCTATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600  
541 CTCTCTCTGAGGCTATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600  
601 ATCACTGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
601 ATCACTGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
661 CATGCTAAATACCAAGTCCCTGGGCTGGGACCATTTGAATCAGTTGAATCTTACAC 720  
661 CATGCTAAATACCAAGTCCCTGGGCTGGGACCATTTGAATCAGTTGAATCTTACAC 720  
721 CTGATATATTTCTGTTATGATGCTGGGTTGAAATCCCATGGCAAGCCATCTTCAAG 780  
721 CTGATATATTTCTGTTATGATGCTGGGTTGAAATCCCATGGCAAGCCATCTTCAAG 780  
781 GTCTCTCTTCAATCTGCAAGCACTATGCTCAGATAGTCTCTTCTGAGCTTTTGG 840  
781 GTCTCTCTTCAATCTGCAAGCACTATGCTCAGATAGTCTCTTCTGAGCTTTTGG 840  
841 TGCCTGCTGATGGCTCTTACCCGCTATGATGATGAGCTATTTGAGCTTCCACAG 900  
841 TGCCTGCTGATGGCTCTTACCCGCTATGATGATGAGCTATTTGAGCTTCCACAG 900  
901 AACGAGCTGCTAGGGGATCCAGATCCAGAGCTAAAGAGAGAGAGAGAGAGAGAG 960  
901 AACGAGCTGCTAGGGGATCCAGATCCAGAGCTAAAGAGAGAGAGAGAGAGAGAG 960  
961 CCGATGCTTTCAGATCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
961 CCGATGCTTTCAGATCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
1021 TCAGCTGCTGATATGCTCTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 TCAGCTGCTGATATGCTCTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1081 CGGAATATCTACAGCTTCTTCCAGAAATGCAATGCAAGAGAAATTTGTTGGGCTC 1140  
1081 CGGAATATCTACAGCTTCTTCCAGAAATGCAATGCAAGAGAAATTTGTTGGGCTC 1140  
1141 ATGAGATATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
1141 ATGAGATATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

1201 ACTGTATAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
1201 ACTGTATAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
1261 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
1321 TTTGACCTATCTGAGATATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1321 TTTGACCTATCTGAGATATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
1381 TTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1441 ACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
1501 CTATTTGAAAGTGAACCTTGCCTTCAAAATTAAGTATGATGATGATGATGATG 1560  
1501 CTATTTGAAAGTGAACCTTGCCTTCAAAATTAAGTATGATGATGATGATGATG 1560  
1561 CACAGTGAAGAGACATGAGACATGACATTTCTAGTCAAGATGAAATATCAATTA 1620  
1561 CACAGTGAAGAGACATGAGACATGACATTTCTAGTCAAGATGAAATATCAATTA 1620  
1621 GAACCTGACCTGGAACCTGCGAGAGCTTACCTCAAGTCACTTCACTTCACTT 1680  
1621 GAACCTGACCTGGAACCTGCGAGAGCTTACCTCAAGTCACTTCACTTCACTT 1680  
1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1741 TGA 1743  
1741 TGA 1743

RESULT 4  
US-10-724-806-23  
; Sequence 23, Application US-10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM  
; APPLICANT: PETERSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:00805  
; CURRENT APPLICATION NUMBER: US/10/724,806  
; PRIOR FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US/09/911,077A  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1743  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1743)  
US-10-724-806-23

Query Match 100.0%; Score 1743; DB 9; Length 1743;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTTTTCCATGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
1 ATGCTTTTCCATGTTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

```

OY 61 GTTGAATATAGGGCTCATGAAAAACCAAAACAGGGCAACCCAGAGAGCGAGTGA 120
DB 61 GTTGAATATAGGGCTCATGAAAAACCAAAACAGGGCAACCCAGAGAGCGAGTGA 120
OY 121 GCCATATAGTGGGGGGCCGTGAATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
DB 121 GCCATATAGTGGGGGGCCGTGAATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
OY 181 ACCGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 ACCGTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
OY 241 GGTCTAGTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 241 GGTCTAGTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
OY 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTAGTATGTAGACCAATCAACAG 360
DB 301 TTTTTCGCAAACTATGCGTTCCAAAGGATATGTAGTATGTAGACCAATCAACAG 360
OY 361 ATCTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 ATCTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
OY 421 TGGGCTGAGCAATTTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
DB 421 TGGGCTGAGCAATTTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
OY 481 GTGAACATATCGGTCTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 540
DB 481 GTGAACATATCGGTCTCTCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 540
OY 541 CTCTACTCTGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 CTCTACTCTGAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 600
OY 601 ATCAGAGTCCCTTTTSCCTGTGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 ATCAGAGTCCCTTTTSCCTGTGATGATGATGATGATGATGATGATGATGATGATG 660
OY 661 CATGCTAAATACAGAGTCCCTGTGATGATGATGATGATGATGATGATGATGATGATG 720
DB 661 CATGCTAAATACAGAGTCCCTGTGATGATGATGATGATGATGATGATGATGATGATG 720
OY 721 CTGTAATATTTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 780
DB 721 CTGTAATATTTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 780
OY 781 GTCTCTCTTCACTCCAGAGTCCCTGTGATGATGATGATGATGATGATGATGATGATG 840
DB 781 GTCTCTCTTCACTCCAGAGTCCCTGTGATGATGATGATGATGATGATGATGATGATG 840
OY 841 TSCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 TSCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
OY 901 AACCAAGTCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 AACCAAGTCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
OY 961 CCGATGTTCTGAGAGTCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 961 CCGATGTTCTGAGAGTCCCTGAGAGATGATGATGATGATGATGATGATGATGATGATG 1020
OY 1021 TCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 TCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
OY 1081 CGGAATATCTACAGCTTTCTTCAAGCAAAATGATGATGATGATGATGATGATGATG 1140
DB 1081 CGGAATATCTACAGCTTTCTTCAAGCAAAATGATGATGATGATGATGATGATGATG 1140

```

```

OY 1141 ATGAGATCATGTGCTTGTGTTGGAGCATGCAACCCATGAGCTTTGTGAGAG 1200
DB 1141 ATGAGATCATGTGCTTGTGTTGGAGCATGCAACCCATGAGCTTTGTGAGAG 1200
OY 1201 ACTGTGTATGGCTCTGTGATCTGAGCTGATGATGATGATGATGATGATGATGATG 1260
DB 1201 ACTGTGTATGGCTCTGTGATCTGAGCTGATGATGATGATGATGATGATGATGATG 1260
OY 1261 CTGCTGTGTACTCTTCAAGAGCAACCACTATGAGGAGCTGCTGATGATGATGATG 1320
DB 1261 CTGCTGTGTACTCTTCAAGAGCAACCACTATGAGGAGCTGCTGATGATGATGATGATG 1320
OY 1321 TTTGAGCTATTCCTGAGATTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
DB 1321 TTTGAGCTATTCCTGAGATTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1380
OY 1381 TTCTACCTGTTTATTTACTCTGACAAAGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 TTCTACCTGTTTATTTACTCTGACAAAGATGATGATGATGATGATGATGATGATGAT 1440
OY 1441 ACTCTCTCAGTGTATACCTCATTTCTTACCAATGATGATGATGATGATGATGATGAT 1500
DB 1441 ACTCTCTCAGTGTATACCTCATTTCTTACCAATGATGATGATGATGATGATGATGAT 1500
OY 1501 CTATTTGAAAGTGAACCTTGCTCCAAATTAATGATGATGATGATGATGATGATGATG 1560
DB 1501 CTATTTGAAAGTGAACCTTGCTCCAAATTAATGATGATGATGATGATGATGATGATG 1560
OY 1561 CACAGTGAAGAGACATGAGCAAGACATTTCTGATGATGATGATGATGATGATGATGAT 1620
DB 1561 CACAGTGAAGAGACATGAGCAAGACATTTCTGATGATGATGATGATGATGATGATGAT 1620
OY 1621 GAACCTTGACCTGTGAACCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1680
DB 1621 GAACCTTGACCTGTGAACCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1680
OY 1681 GAGGCCCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 GAGGCCCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
OY 1741 TGA 1743
DB 1741 TGA 1743

RESULT 5
US-09-911-077A-5
; Sequence 5, Application US/0911077A
; Publication No. US2003011439A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELEY, RANDY D.
; APPLICANT: APPARUNDARUM, SUBRAMANIAM
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA
; FILE REFERENCE: VBEL:008US
; CURRENT APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4904
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; NAME/KEY: CDS
; LOCATION: (224)..(1966)
US-09-911-077A-5

Query Match 92.7%; Score 1616.6; DB 3; Length 4904;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 1664; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

OY 1 ATGCTTTTCAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

```

Db 224 ATGCTTTCATGTAAGAGCTAGTAGGATTATCTGTTCTACCTTTATATTTCTG 283  
Qy 61 GTTGAATATGAGCTGCATGTGAAAAACAACAGCGGCAACCCAGAAAGCGCAGTGA 120  
Db 284 GTTGAATATGAGCTGCATGTGAAAAACAACAGCGGTAATGCAAGAAAGCAGCGAA 343  
Qy 121 GCCATCATATGCGGGGCGCTGACATGTGTTGTTGTTGTTGTTTACATGACAGCC 180  
Db 344 GCCATCATATGCGGGGCGCTGACATGTGTTGTTGTTGTTTACATGACAGCC 403  
Qy 181 ACCTGGTTGAGAGAGCTACATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
Db 404 ACCTGGTTGAG 463  
Qy 241 GGTCTAGCTTGGGCTCATGACCAATGTGATATCTGAGTCTAATTTTAAAGTGGTCTG 300  
Db 464 GGTCTAGCTTGGGCTCATGACCAATGTGATATCTGAGTCTAATTTTAAAGTGGTCTG 523  
Qy 301 TTTTGGGAAACCTATGAGGTTCCAAAGGATATGTGACTATGTATGACCATTCACAAAG 360  
Db 524 TTTTGGGAAACCTATGAGGTTCCAAAGGATATGTGACTATGTATGACCATTCACAAAG 583  
Qy 361 ATCTATGAAAGCGAGTGGGCTGCTCATCCCTGCACTGATGAGAGAGAGAGAGAG 420  
Db 584 ATCTATGAAAGCGAGTGGGCTGCTCATCCCTGCACTGATGAGAGAGAGAGAGAG 643  
Qy 421 TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACCATCAGCGATCATTTGATGTGAT 480  
Db 644 TGGGCTGAGCAATTTTCTGCAATTAAGGGCCACCATCAGCGATCATTTGATGTGAT 703  
Qy 481 GTGAACATATCGGTGATGTTCTCTGCACTATGTCATTTCTTATACCTTAGTGGTGG 540  
Db 704 GTGAACATATCGGTGATGTTCTCTGCACTATGTCATTTCTTATACCTTAGTGGTGG 763  
Qy 541 CTCTACTCTGAGGCAATATCTGATGTTGTCAGCTATTCGATTTTATAGACCTGAG 600  
Db 764 CTCTACTCTGAGGCAATATCTGATGTTGTCAGCTATTCGATTTTATAGACCTGAG 823  
Qy 601 ATCAGTGTCCCTTTTGGCCCTGTACATCTGTGACATCCGACATCGGATTCACAGTGTG 660  
Db 824 ATCAGTGTCCCTTTTGGCCCTGTACATCTGTGACATCCGACATCGGATTCACAGTGTG 883  
Qy 661 CATGCTAATATCCAAAGTCCCTGAGTGGGAAACCATTTGATGAGTGTACACCTG 720  
Db 884 CATGCTAATATCCAAAGTCCCTGAGTGGGAAACCATTTGATGAGTGTACACCTG 943  
Qy 721 CTTGATTAATTTCTGTTATTTGATGCTGAGTGAATCCATGCAAGCCTACTTCCAGAG 780  
Db 944 CTTGATTAATTTCTGTTATTTGATGCTGAGTGAATCCATGCAAGCCTACTTCCAGAG 1003  
Qy 781 GTCTCTCTTCAATCTCAGCAACCTATGCTCAAGTATGCTCTCTCTGAGCCTTTTGGG 840  
Db 1004 GTCTCTCTTCAATCTCAGCAACCTATGCTCAAGTATGCTCTCTCTGAGCCTTTTGGG 1063  
Qy 841 TGCCGTGATAGGCTTACCCGCAATATGAGATGAGATGAGCTTCCACAGATCG 900  
Db 1064 TGCCGTGATAGGCTTACCCGCAATATGAGATGAGATGAGCTTCCACAGATCG 1123  
Qy 901 AACCAAGACTGCTAGCGGATCCAGATCCCAAGACTAAGAGAGAGAGAGAGATTC 960  
Db 1124 AACCAAGACTGCTAGCGGATCCAGATCCCAAGACTAAGAGAGAGAGAGATTC 1183  
Qy 961 CCGATGCTTCTGAGTACTGCTGCTGTGATCATCTCTCTTCTGAGGCTGTGCTGTT 1020  
Db 1184 CCGATGCTTCTGAGTACTGCTGCTGTGATCATCTCTCTTCTGAGGCTGTGCTGTT 1243  
Qy 1021 TCAGTGTGCTGATGCTCAGCTGATCTGATCTGATCTGCTGAGTCTTATGTTGCT 1080  
Db 1244 TCAGTGTGCTGATGCTCAGCTGATCTGATCTGATCTGCTGAGTCTTATGTTGCT 1303  
Qy 1081 CGGAATATCTACAGCTTCTCTGAGCAAAATGATCAGAGAGAGAAATGTGTGGGTC 1140

Db 1304 CGGAATATCTACAGCTTCTCTGAGCAAAATGATCAGAGAGAGAAATGTGTGGGTC 1363  
Qy 1141 ATGAGATCACTGTGCTTGTGTGTGGAGCATCTGCAACAGCCATGCTTGTGACGAG 1200  
Db 1364 ATGAGATCACTGTGCTTGTGTGTGGAGCATCTGCAACAGCCATGCTTGTGACGAG 1423  
Qy 1201 ACTGTGATGAGGCTCTGCTGATCTGAGCTGACCTGTGCTATCATGATCTTCCACAG 1260  
Db 1424 ACTGTGATGAGGCTCTGCTGATCTGAGCTGACCTGTGCTATCATGATCTTCCACAG 1483  
Qy 1261 CTGCTCTGTACTCTTCTCATCAAGAGAACCAACTTATGAGGAGAGTGTGCTTATAT 1320  
Db 1484 CTGCTCTGTACTCTTCTCATCAAGAGAACCAACTTATGAGGAGAGTGTGCTTATAT 1543  
Qy 1321 TTTGACATATCTGAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
Db 1544 TTTGACATATCTGAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1603  
Qy 1381 TTTCAACCTGGTATTAATCTGAGCAAGAAATGTATATCAATCAAGAGTCCCATTTAA 1440  
Db 1604 TTTCAACCTGGTATTAATCTGAGCAAGAAATGTATATCAATCAAGAGTCCCATTTAA 1663  
Qy 1441 ACTCTCTCATAGTATCTCATCTTCTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1500  
Db 1664 ACTCTCTCATAGTATCTCATCTTCTTACCAACATTTGTGTTCTTATCTAGCCAGTAT 1723  
Qy 1501 CTATTTGAAAGTGAACCTTGTCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAG 1560  
Db 1724 CTATTTGAAAGTGAACCTTGTCTCCAAATTTAGATGATTTGATGCTGTGTGCGAAG 1783  
Qy 1561 CACAGTGAAGAGAAATGAGAGAACCACTCTAGTCAGAAATGAAATGAAATTAATTAAT 1620  
Db 1784 CACAGTGAAGAGAAATGAGAGAACCACTCTAGTCAGAAATGAAATGAAATTAATTAAT 1843  
Qy 1621 GAACCTGACCTGTGAAACCTGCGAGAGGCTTAACTTCACTTCACTTCACTTCACT 1680  
Db 1844 GAACCTGACCTGTGAAACCTGCGAGAGGCTTAACTTCACTTCACTTCACTTCACT 1903  
Qy 1681 GAGGCTCTCTTGTATGTTGATTTCCAGTCCGAGAGGCTTGGAGCTGAGATTAATTA 1740  
Db 1904 GAGGCTCTCTTGTATGTTGATTTCCAGTCCGAGAGGCTTGGAGCTGAGATTAATTA 1963  
Qy 1741 TGA 1743  
Db 1964 TGA 1966

RESULT 6  
US-10-724-806-5  
; Sequence 5, Application US/10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:00805  
; CURRENT APPLICATION NUMBER: US/10/724,806  
; PRIOR FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US/09/911,077A  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patencin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 4904  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (224)..(1966)  
US-10-724-806-5  
Query Match 92.7%; Score 1616.6; DB 9; Length 4904;



Query Match	78.9%;	Score 1375;	DB 3;	Length 1743;
Best Local Similarity	86.8%;	Pred. No. 0;		
Matches 1513; Conservative	0;	Mismatches 230;	Indels 0;	Gaps 0;

OY	ATGCCCTTCCAGTNGGGAAGACCTGGTACTCTATTAATCCCTTCTACCTCCCTAATATTTCTG	60
Db	1 ATGGCTTTCCATGTGGAAGAGACTGATAGCTATCATCGTGTCTTCTACCTTCTAATTTTGTCTG	60
OY	61 GTTGGAAATATGCGCTGCATGTGAAAACCAAAAACAGCGGCACACCAGAGAGCGCAGTGA	120
Db	61 GTTGGAAATATGGCGCTGCCTGGAGAAACAAAAACAGTGCACGCGCGAGAAAGACGACGCCAA	120
OY	121 GCCATCATAGTCCGAGGGGCGGTGACATTTGGTTGGTTGGTTGGTTTAAACATGACAGCC	180
Db	121 GCCATCATATGTTGGTGGCGCGAGATATTTGGTTATTTGGTTGGTGGATTTTACATGACACT	180
OY	181 AACTGGGTTGAGAGAGGCTAATCATAGTGGACAGCAGAGAGAGATGTATGGCCAGTTGT	240
Db	181 AACTGGGTCGAGAGGGGTATATCATATGGCACAGCTGAGAGAGTTATGTATACAGTTAT	240
OY	241 GGCTAGCTTGGGGCTCATGCAACCCATTGGATATTTCTGAGTCTAATTTTAAAGTGGTCTG	300
Db	241 GGCTAGCTTGGGGCTCAGGCACCAATTTGGATATTTCTTGAATCTGATTTTAAAGTGGCTG	300
OY	301 TTTTGTGGAAACCTTAGTCGTTCCAAAGGGATATGTACTATGTTAAGCCATTCAACAG	360
Db	301 TTTCTTGGCAAAACCTATCGTTTCAAAAGGGATATGTAGCCATGTATAGACCCGTTTCAGCA	360
OY	361 ATCTATGGAAGCGGATGGGTGGGGGTGCTCTTCAATCCCTGCACTGATGGGAGAGATGTTTC	420
Db	361 ATCTATGGAAGACGATGGGCGAGACTCCGTATTATTCCTGCACTGATGGGAGAAATGTTTC	420
OY	421 TGGGCTGAGCAATTTTCTGTGCATTTAAGGGGCGACCAATCACGCGTATCATTTGATGTGAT	480
Db	421 TGGGCTGAGCAATTTTCTGTCTTTTGGAGACCAACATCACGCGTATCATGATGTGAT	480
OY	481 GTGAACATATGGGTATGTCTGTGCACTCAATTTGCACTTTTAACTTAACTTATGGGTGGG	540
Db	481 ATGCACATTTCTGTATCATCTGTGCACCTCATGTGCACCTGTGACACTGGTGGAGGG	540
OY	541 CTTAATCTGTGGCATATATCATGTATGTTGTCCAGCTATTTCTGCATTTTAAAGACTGGG	600
Db	541 CTTAATCTGTGGCCTTACATGATGTCTGTAGCTCTTTTGCATTTTGTAGGGCTGGG	600
OY	601 ATCAGTGTCCCTTTTGGCCGTGCACATCCTGCAGACACCGCATGTGGATTTCAAGCTGTG	660
Db	601 ATCAGCGTCCCTTTTGCATGTGCATCTCCTGCAGTCCGACATGTGGGATTTCACTGTGTG	660
OY	661 CATGTAAATATACAGAGTCCCTGTGGGAAACCATTTGATAGTTGAAGTCTACACTGG	720
Db	661 CATGCAATATACAAAACCCGTGGCTGGGAACGTGTGACTCATGTGAGTCTACTCTTGG	720
OY	721 CTTGATATATTTTCTGTATATGATGCTGGGTGGAAATCCATGGCAAGCCTACTTCCAGAG	780
Db	721 CTTGATATATTTTCTGTATGTATGTGTGGGTGGAAATCCCATGGCAAGCATCTTCCAGGG	780
OY	781 GTCTCTCTTATCTCTACGCCACCTATCCTCAGTACTGTCTTCTTGGACAGCTTTTGGG	840
Db	781 GTTCTCTCTTCTCTCTACGCCACCTATCTCAAGTGTCTTCTTCTTGGACAGCTTTTGGG	840
OY	841 TGCCGTGGAGATGGCTTACCCGCCATATGCAATAGCAAGCATTTTGGAGCTTCCACAGACTGG	900
Db	841 TGCCGTGGAGATGGCCATCCACGCCATCTCATTTGGGGCCATTTGGAGCAATCAACAGACTGG	900
OY	901 AACCAAGACTGCGCTACGGGTATTCGAATCCCAAGACTAAGGAGAGAGCAGATGATTTCTC	960
Db	901 AACCAAGACTGCAATATAGGGCTTCCAGATCCCAAGACTTACAGAAAGGAGCAGATGATTTT	960
OY	961 CCGATCGTTTCTGCACTACTCTGCCCTGTGTACATCTCTTCTTGGGCTGTGTCTGTT	1020
Db	961 CCAATGTGTCGCAAGTATCTGCGCCCTGTGTATATTTCTTCTTGTGTGTGGTGTGCACTT	1020

OY	1021	TCACCTGCTGTCAGTCTCTCAGCTGACCTGCTCCATCTCGTCGGGAGATTCTAGTTGGCT	1080
Db	1021	TCGCTGCTGTTATGTATCAGCAGATTTCTTCCATCTTGTCAGCAATTTCAATTTTGCA	1080
OY	1081	CGGAATATCTACAGCTTTCTCTCAGCAAAATGATCAOAGCAAGAAATTGTGSGGTC	1140
Db	1081	CGGAACATCTACAGCTTTCTCTCAGCAAAATGCTTCGCAAAAGAAATGTTGGGTT	1140
OY	1141	ATGAGGAATCACTGAGCTTGTTGGAGCAATCTGSCAACAGCATGAGCTTTGTGACGAAG	1200
Db	1141	ATGCGGAATCAACAGTGTGTGGTTGGAGCAATCTGCAACAGCATGAGCTTGCTGACGA	1200
OY	1201	ACTGTGTATGAGGCTCTGGTACCTGATCTGAGCTTGACCTTGTCTPACATCATCTTCCACAG	1260
Db	1201	ACTGTGTATGAGGCTCTGGTACCTGATCTGAGCTTGACCTTGTTAACATCGTTATCTTCCCCAG	1260
OY	1261	CTGCTCTGTGACTCTTCATCAAAAGAACCAACCTTAATGGGAGAGTGTGCTGTTATATT	1320
Db	1261	CTGCTCTGTGACTCTTGTGTTAAAGGAAACCAACCTTAATGGGAGCTGAGGAGTTATGTT	1320
OY	1321	TTTGACATTAATCCGAGAAATTAATCGAGAGAGGCATATCTAATCTTGACAGCCCTTAATC	1380
Db	1321	TTGTGGCTCTTCCGAGAAATPACTGAGAGGAGGCATATCTGTAATCTTCAGCCCTTGATC	1380
OY	1381	TTCTACCTGCTTAATCTGTGACAGAAATGTAATATACATCAGAGGTTCCATTTAA	1440
Db	1381	TTCTACCTGCTTAATCCCTGATGATTAATGTAATATATATCAGAAATTTCCATTTAA	1440
OY	1441	ACTCTCCAGATGCTTAATCTTTAACCAACATTTGTGTTCTTPTACTAGCCAATAT	1500
Db	1441	ACACTGTGCATGTTACATCTTTTAAACCAACATTTGCATCTCTTAATCTTAAGCCAATAT	1500
OY	1501	CTATTGTAAGTGAACCTTGCTCCCAAAATTTAGATGTAATGATGCTGTTGTGCAAGG	1560
Db	1501	CTATTGTAAGTGAACCTTGCTCCCACTTAATGATGTAATGATGCTGTTGTGCAAGA	1560
OY	1561	CACGTGAAGAACATGAGCAACATCTTGTGGAATGAAATTAATCAATTTAAT	1620
Db	1561	CACAGTGAAGAAACATGAGATTAACAAATTTGTCAAAATGAAATTAATTAATGAT	1620
OY	1621	GAACCTGCACCTGTGAACCTCGGACAGAGCTTAACCTCTAGTTCAACTTTCACCAATAAG	1680
Db	1621	GAACTTGCACCTGTGAAGCCACGACAGATGATCCTCAGCTCAACTTTCACCAATAAA	1680
OY	1681	GAGGCTCTTCCATTGATTGTAATCCAGTCCAGAAAGGCTGTGGAGCTGAAGATTAATTACAG	1740
Db	1741	TGA 1743	
OY	1741	TGA 1743	
Db	1741	TGA 1743	
RESULT 8			
US-10-724-806-1			
Sequence 1, Application US/10724806			
Publication No. US2004024838A1			
GENERAL INFORMATION:			
APPLICANT: BLARELY, RANDY D.			
APPLICANT: APPARUSUNDARAM, SUBRAMANIAM			
TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA			
FILE REFERENCE: VBLT:008US			
CURRENT APPLICATION NUMBER: US/10/724,806			
CURRENT FILING DATE: 2003-12-01			
PRIOR APPLICATION NUMBER: US/09/911,077A			
PRIOR FILING DATE: 2001-07-23			
NUMBER OF SEQ ID NOS: 27			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 1			
LENGTH: 1743			
TYPE: DNA			
ORGANISM: Homo sapiens			

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1743)
US-10-724-806-1

```

Query Match	78.9%	Score 1375;	DB 9;	Length 1743;
Best Local Similarity	86.8%;	Pred. No. 0;		
Matches 1513;	Conservative	0;	Mismatches 230;	Indels 0;
			Gaps	0

Qy	1	ATGCGCTTTCACATGAGAAAGACCTGGATGATATATCTCTCTACCCCTCTATATATCTG	60
Dp	1	ATGGCTTTCACATGAGAAAGACTGATATCATCTGTCCTACCTCTAAATTTGCTG	60
Qy	61	GTTGGAAATATGAGGCTGCATGAGAAAACCAAAAACAGCGGCACCCAGAGGCCACTGMA	120
Dp	61	GTTGGAAATATGAGGCTGCCTCGAGAAACCAAAAACAGTGGACGCCAGAGGCCACCGAA	120
Qy	121	GCCATCATATAGTGGAGGAGCGGTGACATTTGGTTTGTTGGTGGTGGTTTACATGACAGCC	180
Dp	121	GCCATCATATAGTTGGTGGCCGAGATATTTGGTTTATTTGGTTGGTGGATTTACATGACAGCT	180
Qy	181	ACCTGGGTTGAGAGAGGCTACATCAATGGGACAGCAAGAACAGTGTATGGGCCAGGTTGT	240
Dp	181	ACCTGGGTGGAGAGGGGTATATCAATGGCACAGCTGAAGCACTTATATGACCAAGGTTAT	240
Qy	241	GCTCTAGCTTGGGCTCATGCACCCATTGGATATTTCTCTGAGTCTAAATTTTAAAGTGGCTG	300
Dp	241	GGCCTTAGCTTGGGCTCAGGACCAATTTGATATTTCTTTAGTCTGATTTTAAAGTGGCTG	300
Qy	301	TTTTTTGGGAACCTATGCGTTCCAAAGGATATGACATATGTTATGACCCCATCAACAG	360
Dp	301	TTCTTTGCAAAACCTATGCGTTCAAAAGGGATATGACATGTTAAGACCCGTTTCAGCAA	360
Qy	361	ATCTATGGAAGACGCGATGSGTGGGCTGCTCTTCATCCCTGCACCTGATGGAGAGATGTTTC	420
Dp	361	ATCTATGGAAGAACGCAATGGGCGGACCTCGTATTATCTGTCACCTGATGGGAGAAATGTTTC	420
Qy	421	TGGGCTGACGCAATTTTCTCTGCAATTAAGGAGGCCACATACAGCGTATCATTTGATG	480
Dp	421	TGGGCTGACGCAATTTTCTCTGTTTGGGAGGCCACCATACAGCGTATCATTCAGATG	480
Qy	481	GTTGACATATCGGTCATTTGCTCTGACATCTTGCATTTGCTTATATACCTAGTGGGTGG	540
Dp	481	ATGCAATTTCTGTATCATCTCTGTGACATCATTTGCCACTGTGACACACGTTGGGAGGG	540
Qy	541	CTCTACTCTGTGGCAATATCTGATGTTGTGTCAGACATTTGCAATTTTATATAGGACTGTG	600
Dp	541	CTCTATTTCTGTGGCTCACACTGATGTGTTGCTTACGCTCTTTTGCAATTTTGTAGGGGCTGTG	600
Qy	601	ATCAGTGTCCCTTTTGGCCCTGTACACATCTCTGCAGTACCGGACATCGGATTCACAGCTGTG	660
Dp	601	ATCAGGTCCTCTTTTGTCATTTGTTCACATCTCTGCAGTGGCAGACATCGGGTTCACTGCTGTG	660
Qy	661	CATGCTAAATACGAGTCCCTGGCTGGGAAACCATTTGATCAGTTGAAGTCTACACCTGG	720
Dp	661	CATGCAAAATACCAAAAGCCGTGGCTGGGAAACGTGACTCATCTGAAGCTCATCTTTGG	720
Qy	721	CTGTGTAATTTCTGTATTTGATATGCTGGGGTGGAAATCCCATGGGAAGCTCATCTCCAGAG	780
Dp	721	CTGTGTAAGTTTCTGTGTTGTTGATGCTGGGGTGGAAATCCCATGGGACGATATCTTCAGAG	780
Qy	781	GTCCTCTCTTCATCTCTCAGCCACCTATGCTCAGAGTCTGTCTTCTCTGGACGCTTTTGGG	840
Dp	781	GTTCTCTCTTCTCTCAGCCACCTATGCTCAAGTGTGTCTTCTCTGGACGCTTTTGGG	840
Qy	841	TGCTCTGTATGGCTTACCCCGCATATGATATGAGACTATTTGAGCTTTCACAGACTGG	900
Dp	841	TGCTCTGTATGGCCATCCGACATATCTATTTGGGGCAATTTGAGGATTAACAAGACTGG	900
Qy	901	AACCAAGACTGCTACGAGGTATCCAGATCCCAAGCTAAAGAGGAGAGACATGATCTC	960
Dp	901	AACCAAGACTGCTATGGGCTTCCAGATCCCAAGCTACAGAAAGGACAGACATGATTTTA	960

OY	961	CCGATTCGTCGAGTACCTCTGCCCTGTGATCTCTCTTTGGGCTTGGTCTGT	1020
Db	961	CCATTTGTTCCGAGTATCTCTGCCCTGTGATATTTCTTTCTTTGGCTTGGTCA	1020
OY	1021	TCAGCTGCTGATGTCCTCAGTGACTCGTCCATCCTGTGGCGAGTTCTATGTTGCT	1080
Db	1021	TCCTCTGCTGTATGTCATCAGCAGATTCCTTCCATCTGTGCAGAAATTCATGTTTGA	1080
OY	1081	CGGAATATCTACAGAGCTTTCCTTCAGCAAAATGATCATCAGACAGAAATTGTTGGGCT	1140
Db	1081	CGGAATATCTACAGAGCTTTCCTTCAGCAAAATGATCATCAGACAGAAATTGTTGGGCT	1140
OY	1141	ATGAGGATCACTGTCCTGTGTGTGGAGCATCTGCAACAGCATGCTTTGTCTGACGAA	1200
Db	1141	ATGAGGATCACTGTCCTGTGTGTGGAGCATCTGCAACAGCATGCTTTGTCTGACGAA	1200
OY	1201	ACTGTGTATGGGCTCTGTGACTGTGACCTGTGACCTTGTCTATCATCATCTTCCACAG	1260
Db	1201	ACTGTGTATGGGCTCTGTGACTGTGACCTGTGACCTTGTCTATCATCATCTTCCACAG	1260
OY	1261	CTGCTCTGTGACTCTTCATCAAGAGAACCAACTATGGGGCAGTTGCTGTTATAT	1320
Db	1261	CTGCTCTGTGACTCTTCATCAAGAGAACCAACTATGGGGCAGTTGCTGTTATAT	1320
OY	1321	TTTGGACTATTCCTGAGAAATTACTGGAGGAGGCCATATCTTAACTTGACGCTTAACT	1380
Db	1321	TTTGGACTATTCCTGAGAAATTACTGGAGGAGGCCATATCTTAACTTGACGCTTAACT	1380
OY	1381	TTTCAACCTGGTTATTAATCTGTGACAAAGATGTTATATACATCGAGGTTCCCATTTAA	1440
Db	1381	TTTCAACCTGGTTATTAATCTGTGACAAAGATGTTATATATATCAAGAAATTTCCATTTAA	1440
OY	1441	ACTCTCTCCAGTGTACTCATCTCTTAAACCAACTTTGTGTTCTTAACTTCAAGCCAGTAT	1500
Db	1441	ACACTTGCATGTTACATCATCTTAAACCAACTTTGCATCTCTTATCTAGCCAGATAT	1500
OY	1501	CTATTTGAAAGTGAACCTTGCTCCCAAAATTAATGATGATTTGATGCTGTGTGCAAG	1560
Db	1501	CTATTTGAAAGTGAACCTTGCTCCCACTAAATTAATGATGATTTGATGCTGTGTGCAAG	1560
OY	1561	CACAGTGAAGGAATGTGACAAACCACTTCTACGAAATGAAATATCAAAATTAAT	1620
Db	1561	CACAGTGAAGGAACATGTGATGAACCAATTTCTGTCAAAATGAAATATTTAAATTAAT	1620
OY	1621	GAACTTGACCTGTGAACCTCGGCAGAGCCCTCAAGTCTCAACTTTCAACCAATTAAG	1680
Db	1621	GAACTTGACCTGTGTGAAGCCACGACAGAGCATGACCTCGACCTTACACCAATTAAG	1680
OY	1681	GAGGCGCTCTCTGATGTTGATTTCCAGTCCAGAAAGGCTCTGGGACTGAAGATTAATTAAG	1740
Db	1681	GAGGCGCTCTCTGATGTTGATTTCCAGTCCAGAAAGGCTCTGGGACTGAAGATTAATTAAG	1740
OY	1741	TGA TGA TGA	1743
Db	1741	TGA TGA TGA	1743

```

RESULT 9
US-09-911-077A-9
; Sequence 9, Application US/09911077A
; Publication No. US20030114399A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARETUNABARM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNN
; FILE REFERENCE: VBLT.008US
; CURRENT APPLICATION NUMBER: US/09/911.077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1813

```



TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (19)..(1761)  
US-09-911-077A-9

Query Match 78.9%; Score 1375; DB 3; Length 1813;  
Best Local Similarity 86.8%; Pred. No. 0;  
Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```
1 ATGCTTCCATGGAAGGAGCTGGATGCTATTCCTTACCTCCTTATTTTCG 60
19 ATGGCTTCCATGGAAGGAGCTGGATGCTATTCCTTACCTCCTTATTTTCG 78
61 GTTGAATATGAGGCTGATGGAACCAAAACAGCGGCAACCCAGAGGCGAGTGA 120
79 GTTGAATATGAGGCTGATGGAACCAAAACAGCGGCAACCCAGAGGCGAGTGA 138
121 GCCATCATATGCTGGGGCCGCTGACATGCTGTTGTTGTTGTTTACATGACAGCC 180
139 GCCATCATATGCTGGGGCCGCTGACATGCTGTTGTTGTTGTTTACATGACAGCT 198
181 ACCGTGGTGGAGGAGGCTACATCATGAGGAGAGAGAGAGAGAGAGAGAGTTC 240
199 ACCGTGGTGGAGGAGGCTACATCATGAGGAGAGAGAGAGAGAGAGAGAGTTC 258
241 GGTCTAGCTGGGCTCATGACACCATGGAATTCCTGAGTCTAATTTTAAAGTGTCTG 300
259 GGGCTAGCTGGGCTCATGACACCATGGAATTCCTGAGTCTAATTTTAAAGTGTCTG 318
301 TTTTGGGAAACCTATGCTTCCAAAGGATATGACTATGTTAGACCATTTCAAACAG 360
319 TTTTGGGAAACCTATGCTTCCAAAGGATATGACTATGTTAGACCATTTCAAACAG 378
361 ATCTATGAAAGCGATGGGTGGGCTCTTATCCCGGACATGAGGAGAGAGATTTTC 420
379 ATCTATGAAAGCGATGGGTGGGCTCTTATCCCGGACATGAGGAGAGAGATTTTC 438
421 TGGGCTGACGCAATTTTCTGTCATTAAGGGGCCACCATCAGCGTATCATTTGATGAT 480
439 TGGGCTGACGCAATTTTCTGTCATTAAGGGGCCACCATCAGCGTATCATTTGATGAT 498
481 GTGAACATATCGGTATGCTCTGTCATTTGCCATTTCTTAAACCTATGGGTGGG 540
499 ATGCAATTTTCTGTCATTTCTGTCATTTGCCATTTCTTAAACCTATGGGTGGG 558
541 CTCTACTCTGTCATTTCTGTCATTTGCCATTTCTTAAACCTATGGGTGGG 600
559 CTCTACTCTGTCATTTCTGTCATTTGCCATTTCTTAAACCTATGGGTGGG 618
601 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTCATTTGCCATTTCTTAAACCTATGG 660
619 ATCAGTGTCCCTTTTGGCCCTGTCACATCTGTCATTTGCCATTTCTTAAACCTATGG 678
661 CATGCTAAATACCAAGTCCCTGCTGGGAAACCATTTGAATGAGTGAAGTCTACACTGG 720
679 CATGCTAAATACCAAGTCCCTGCTGGGAAACCATTTGAATGAGTGAAGTCTACACTGG 738
721 CTGATTAATTTTCTGTTATGATGCTGGGATGGAATCCCATGCAAGCCATTTCCAGAGG 780
739 CTGATTAATTTTCTGTTATGATGCTGGGATGGAATCCCATGCAAGCCATTTCCAGAGG 798
781 GTCTCTCTTCACTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGG 840
799 GTCTCTCTTCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGG 858
841 TGCCGTGATGAGGCTCTTACCCGCGCATATGAGAGATTTGAGCTTCCACACTGG 900
859 TGCCGTGATGAGGCTCTTACCCGCGCATATGAGAGATTTGAGCTTCCACACTGG 918
901 AACCAAGTGTCTTACCCGCGCATATGAGAGATTTGAGCTTCCACACTGG 960
```

```
919 AACCAAGTGTCTTACCCGCGCATATGAGAGATTTGAGCTTCCACACTGG 978
961 CGGATGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1020
979 CCAATTTGTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1038
1021 TCAGTGTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1080
1039 TCAGTGTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1098
1081 CGGATGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1140
1099 CGGATGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1158
1141 ATGAGGATGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1200
1159 ATGAGGATGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1218
1201 ACTGTGTATGGGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1260
1219 ACTGTGTATGGGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1278
1261 CTGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1320
1279 CTGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1338
1321 TTTGAGCTATTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1380
1339 TCTGAGCTATTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1398
1381 TTTGAGCTATTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1440
1399 TTTGAGCTATTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1458
1441 ACTGTGTATGGGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1500
1459 ACTGTGTATGGGCTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1518
1501 CTATTTGAAAGTGAACCTTTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1560
1519 CTATTTGAAAGTGAACCTTTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1578
1561 CACAGTGAAGGATGCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1620
1579 CACAGTGAAGGATGCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1638
1621 GAAGTGAAGGATGCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1680
1639 GAAGTGAAGGATGCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1698
1681 GAGGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1740
1699 GAGGCTTCTGTCAGCAACCTATGCTGAGGATGCTGCTTCTGTCAGCTTTTGGGCTGTG 1758
1741 TGA 1743
1759 TGA 1761
```

RESULT 10  
US-10-724-806-9  
; Sequence 9, Application US/10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELY, RANDY D.  
; APPLICANT: APPARUSUNDARAM, SUBRAMANIAM  
; APPLICANT: FERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA  
; FILE REFERENCE: VBLF:00805  
; CURRENT APPLICATION NUMBER: US/10/724, 806  
; PRIOR APPLICATION NUMBER: 2003-12-01  
; PRIOR FILING DATE: 2001-07-23



NUMBER OF SEQ ID NOS: 27  
 SOFTWARE: Patentin Ver. 2.1  
 SEQ ID NO: 9  
 LENGTH: 1813  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (19) .. (1761)  
 US-10-724-806-9

Query Match 78.9%; Score 1375; DB 9; Length 1813;  
 Best Local Similarity 86.8%; Pred. No. 0;  
 Matches 1513; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

```

QY 1 ATGCCCTTTCATGTGGAAGGACTGTAAGTATATCTCTTCACTCTCTATATTTCTG 60
DB 19 ATGGCTTTTCATGTGGAAGGACTGTAAGTATATCTCTTCACTCTCTATATTTCTG 78
QY 61 GTTGAATATGGGCTCATGAAAAACAAAACAGGCGCAACCCAGAGAGCGAGTGA 120
DB 79 GTTGAATATGGGCTCATGAAAAACAAAACAGGCGCAACCCAGAGAGCGAGTGA 138
QY 121 GCCATCATAGTCGGGGCCGTGACATTTGTTGTTGTTGTTGTTTACCATGACAGCC 180
DB 139 GCCATCATAGTTGGTGGCCGAGATATTTGTTTATTTGTTGTTGTTTACCATGACAGCT 198
QY 181 ACCGTGGTGGAGAGGCTCATCATATGAGACAGCAAGAGCATGTATGGCCAGTTGT 240
DB 199 ACCGTGGTGGAGAGGCTCATCATATGAGACAGCAAGAGCATGTATGATACAGTTAT 258
QY 241 GGCTAGCTTGGGCTCATGACCCCATTTGATTTCTGAGTCTATTTTAAAGTGGCTG 300
DB 259 GGCTAGCTTGGGCTCATGACCCCATTTGATTTCTGAGTCTATTTTAAAGTGGCTG 318
QY 301 TTTTTCGCAACCTATGCGTTCCAAAGGATATGATATGTTAGACCCATTCAAACAG 360
DB 319 TTTTTCGCAACCTATGCGTTCCAAAGGATATGATATGTTAGACCCATTCAGCAA 378
QY 361 ATCTATGAAAGGCGATGGGTGGGCTGCTCTTCTATCCCTGCACTGATGGAGAGATGTT 420
DB 379 ATCTATGAAAGGCGATGGGTGGGCTGCTCTTCTATCCCTGCACTGATGGAGAGATGTT 438
QY 421 TGGGCTGACGAAATTTCTCTGATTAAGGGGCCACCATGAGGCTATGATGTTGAT 480
DB 439 TGGGCTGACGAAATTTCTCTGATTAAGGGGCCACCATGAGGCTATGATGTTGAT 498
QY 481 GTGAACATATCGGTCTATGCTCTGCACTATTCATTTCTTTATACCTAGTGGGTGG 540
DB 499 ATGCAATATTTCTGATCTCTGCACTATTCATTTGCACTCTGTAACACCTGGTGGAGGG 558
QY 541 CTCTACTCTGTGGCATATCTGATGTTGTCAGCTATTTCTGCAATTTTATAGAGCTGTG 600
DB 559 CTCTATTTCTGTGGCTTACACTGATGTCGTTCACTCTTTTGAAGGGCTGTG 618
QY 601 ATCAGTGCCTTTTGGCCGTGACATCTGACAGTACAGCAATCGGATTTCAAGCTGTG 660
DB 619 ATCAGTGCCTTTTGGCCGTGACATCTGACAGTACAGCAATCGGATTTCAAGCTGTG 678
QY 661 CATGTAAATACAGAGTCCCTGGCTGGAGAACCATTTGAATCAAGTTCAACCTGG 720
DB 679 CATGTAAATACAGAGTCCCTGGCTGGAGAACCATTTGAATCAAGTTCAACCTGG 738
QY 721 CTGTAAATATTTCTGTATATGATGCTGGTGAATCCCATGGAACCTATCTTCAGAG 780
DB 739 CTGTAAATATTTCTGTATATGATGCTGGTGAATCCCATGGAACCTATCTTCAGAG 798
QY 781 GTCTCTCTTCACTCTCAGCAGCACTATGCTCAGGTACTGTCTTCTGAGAGCTTTGGG 840
DB 799 GTCTCTCTTCTCTCAGCAGCACTATGCTCAGGTACTGTCTTCTGAGAGCTTTGGG 858
QY 841 TGCCTGTGATGGCTTACCCGCAATGATGAGACTATTTGAGCTTCCACAGACTGG 900

```

```

DB 859 TGCCTGTGATGGCCATCCCAAGCACTACTATTTGGGGCCATTGGAGCATCAACAGACTGG 918
QY 901 AACCCAGACTGCTTAAGGGATATCCAGATGCCAAGACTAAAGAGAACAGCATGATCTC 960
DB 919 AACCCAGACTGCTTAAGGGATATCCAGATGCCAAGACTAAAGAGAACAGCATGATTTTA 978
QY 961 CCGATGCTCTGAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 979 CCAATGTTCTGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038
QY 1021 TCAAGCTGTGATGCTCTGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1039 TCGCTGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
QY 1081 CGGAATATCTACAGCTTCTCTGAGCAAAATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1099 CGGAATATCTACAGCTTCTCTGAGCAAAATGATGATGATGATGATGATGATGATGATGAT 1158
QY 1141 ATGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1159 ATGAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1218
QY 1201 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1219 ACTGTGTATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1278
QY 1261 CTGCTGTGTATCTCTCTTCAAAAGAACCACTATGAGGCACTGCTGCTGCTGCTGCTGCT 1320
DB 1279 CTGCTGTGTATCTCTCTTCAAAAGAACCACTATGAGGCACTGCTGCTGCTGCTGCTGCT 1338
QY 1321 TTTGACATATCTCTGAGATTAATCTGAGAGAGAGCCATATCTATCTGAGAGCCCTTAATC 1380
DB 1339 TCTGCTCTCTCTGAGATTAATCTGAGAGAGAGCCATATCTGATCTGATCTGAGCCCTGATC 1398
QY 1381 TTTTACCTGCTTATTAATCTGAGCAAAAGTGTATATCAATCAAGAGTTCCCATTTAAA 1440
DB 1399 TTTTACCTGCTTATTAATCTGAGCAAAAGTGTATATATCAAGAAATTTCCATTTAAA 1458
QY 1441 ACTCTCTCATAGTTATCCCTGATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1500
DB 1459 ACACTCTCATAGTTATCCCTGATTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1518
QY 1501 CTATTTGAAAGTGAACCTTGGCTCCAAATATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1519 CTATTTGAAAGTGAACCTTGGCTCCAAATATGATGATGATGATGATGATGATGATGATGAT 1578
QY 1561 CACAGTGAAGAGAACATGAGCAAGACCAATTTAGTCAGAAATGAATCAAAATTAAT 1620
DB 1579 CACAGTGAAGAGAACATGAGTAAAGCAATTTAGTCAGAAATGAATCAAAATTAAT 1638
QY 1621 GAACCTGCACTGTGAAACCTTGGGCAAGGCTTAACCTGATGATGATGATGATGATGATGAT 1680
DB 1639 GAACCTGCACTGTGAAAGCAAGACAGACATGATGATGATGATGATGATGATGATGATGAT 1698
QY 1681 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1699 GAGGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
QY 1741 TGA 1743
DB 1759 TGA 1761

```

## RESULT 11

US-11-005-030-1  
 ; Sequence 1, Application US/11005030  
 ; Publication No. US20050255495A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aerssens, Jeroen  
 ; APPLICANT: Athanasios, Maria  
 ; APPLICANT: Brain, Carlos  
 ; APPLICANT: Cohen, Nadine  
 ; APPLICANT: Dain, Bradley

```

; APPLICANT: Denton, R. Rex
; APPLICANT: Judson, Richard S.
; APPLICANT: Ozdemir, Vural
; APPLICANT: Reed, Carol R.
; TITLE OF INVENTION: SLC5A7 Genetic Markers Associated with Age of Onset of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2300.0090001
; CURRENT APPLICATION NUMBER: US/11/005,030
; CURRENT FILING DATE: 2004-12-07
; PRIOR APPLICATION NUMBER: US 60/529,999
; PRIOR FILING DATE: 2003-12-15
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent version 3.3
; SEQ ID NO 1
; LENGTH: 26323
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1973)..(1973)
; OTHER INFORMATION: n is 'g' or 'a'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2010)..(2010)
; OTHER INFORMATION: n is 't' or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2273)..(2273)
; OTHER INFORMATION: n is 'a' or 'g'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2417)..(2417)
; OTHER INFORMATION: n is 't' or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7671)..(7671)
; OTHER INFORMATION: n is 'g' or 't'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16520)..(16520)
; OTHER INFORMATION: n is 'c' or 't'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16546)..(16546)
; OTHER INFORMATION: n is 'g' or 'a'
; US-11-005-030-1

Query Match      28.8%; Score 502.8; DB 15; Length 26323;
Best Local Similarity 87.1%; Pred. No. 1.3e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AAATGATCATGACGAAGAAATTTGTGGTTCATGAGATCACTGCTTGTGTCGAGC 1169
DB 24690 ACAGGCTTCGACGAAGAAATCGTTTGGGTTATGCAATCAAGTGTGTTGGTGAAC 24749
QY 1170 ATCTGCAACGACATGCTTGTGCTGACGAAGCTGTGTATGGGCTTGTACTGAGCTC 1229
DB 24750 ATCTGCAACGACATGCTTGTGCTGACGAAGCTGTGTATGGGCTTGTACTGAGCTC 24809
QY 1230 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGACTCTTCTTCAAAAGAAC 1289
DB 24810 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGACTCTTCTTCAAAAGAAC 24869
QY 1290 CAACACTATGAGGAGCTGTGCTGTTATATTTTGGACTATTTCTTGAGATTAAGAGG 1349
DB 24870 CAACACTATGAGGAGCTGTGCTGTTATATTTTGGACTATTTCTTGAGATTAAGAGG 24929
QY 1350 AGAGCCATATCTATCTTGTGACGCCCTTATCTTTTACCTGTGTTATTAATCTGACAGAA 1409
DB 24930 AGAGCCATATCTATCTTGTGACGCCCTTATCTTTTACCTGTGTTATTAATCTGACAGAA 24989
QY 1410 TGGTATATACATCAGAGGTTCCCATTTAAACTCTCTCCATGTTACTCATCTTCTTTAC 1469
DB 1410 TGGTATATACATCAGAGGTTCCCATTTAAACTCTCTCCATGTTACTCATCTTCTTTAC 1469
```

```

DB 24990 TGGTATATATATCAGAAATTTCCATTTAAACACTTCCATGGTTACTCATCTTTAC 25049
QY 1470 CAACATTTGTGTTCTTATCTTACGCAAGATCTATTTTGAAGTGAACCTTGCTCCAAA 1529
DB 25050 CAACATTTGATCTCTATCTTACGCAAGATCTATTTTGAAGTGAACCTTGCTCCAAA 25109
QY 1530 ATTGATGTTATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGACAGACCAT 1589
DB 25110 ATTGATGTTATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGATTAAGACAT 25169
QY 1590 TCTAGTCAGAAATGAATAATCAATTAATGAAGAACTTGCACCTGTGAACCTCGGACAG 1649
DB 25170 TCTTGTCAAAAATGAAAATTAATTAATTAATGAAGAACTTGCACCTGTGAAGCCACGACAG 25229
QY 1650 CCTAACCTCAGTTCACTTTACCAATTAAGAGGAGGCTCTTGATGTTGATTCAGTCC 1709
DB 25230 CATGACCTCAGCTCACTTCACTTCAACCAATTAAGAGGAGGCTCTTGATGTTGATTCAGTCC 25289
QY 1710 GGAGGGCTCTGGAGCTGAAGATTAATTAATGA 1743
DB 25290 AGAAGGCTCTGGAGCTGAAGATTAATTAATGA 25323

RESULT 12
US-09-911-077A-19
; Sequence 19, Application US/09911077A
; Publication No. US20030114599A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARISUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:00805
; CURRENT APPLICATION NUMBER: US/09/911,077A
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 19
; LENGTH: 119040
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (2347)..(90873)
; OTHER INFORMATION: N = A, C, G or T/U
; US-09-911-077A-19

Query Match      28.8%; Score 502.8; DB 3; Length 119040;
Best Local Similarity 87.1%; Pred. No. 3.5e-148;
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AAATGATCATGACGAAGAAATTTGTGGTTCATGAGATCACTGCTTGTGTCGAGC 1169
DB 30755 ACAGGCTTCGACGAAGAAATCGTTTGGGTTATGCAATCAAGTGTGTTGGTGAAC 30814
QY 1170 ATCTGCAACGACATGCTTGTGCTGACGAAGCTGTGTATGGGCTTGTACTGAGCTC 1229
DB 30815 ATCTGCAACGACATGCTTGTGCTGACGAAGCTGTGTATGGGCTTGTACTGAGCTC 30874
QY 1230 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGACTCTTCTTCAAAAGAAC 1289
DB 30875 TGACCTGTCTACATCATCTTCCACAGCTGCTGTGACTCTTCTTCAAAAGAAC 30934
QY 1290 CAACACTATGAGGAGCTGTGCTGTTATATTTTGGACTATTTCTTGAGATTAAGAGG 1349
DB 30935 CAACACTATGAGGAGCTGTGCTGTTATATTTTGGACTATTTCTTGAGATTAAGAGG 30994
QY 1350 AGAGCCATATCTATCTTGTGACGCCCTTATCTTTTACCTGTGTTATTAATCTGACAGAA 1409
DB 30995 AGAGCCATATCTATCTTGTGACGCCCTTATCTTTTACCTGTGTTATTAATCTGACAGAA 31054
QY 1410 TGGTATATACATCAGAGGTTCCCATTTAAACTCTCTCCATGTTACTCATCTTCTTTAC 1469
DB 1410 TGGTATATACATCAGAGGTTCCCATTTAAACTCTCTCCATGTTACTCATCTTCTTTAC 1469
```

Db 31055 TGGTATATATATACAGAAATTTCCATTAAACACTTGCATGGTTACATCATTTCTTAAC 31114  
QY 1470 CAACATTGTGTTCTTATCTAGCAAGATATCTATTGGAAGGAACTTGGCTCCAA 1529  
Db 31115 CAACATTGTGATCTCTTATCTAGCAAGATATCTATTGGAAGGAACTTGGCTCCAA 31174  
QY 1530 ATTAGATGATTTGATGCTGTGTGCGAAGGCAAGTGAAGAGAACATGACAAGCCAT 1589  
Db 31175 ATTAGATGATTTGATGCTGTGTGCGAAGGCAAGTGAAGAGAACATGACAAGCCAT 31234  
QY 1590 TCTAGTCAGAAATGAAATATCAAAATTAATGAACCTTGCACCTGTGAACCTGGCAGAG 1649  
Db 31235 TCTGTCAAAAAGAAATTAATTAATGATGAACCTTGCACCTGTGAAGCCAGACAGAG 31294  
QY 1650 CCTAACCCCTCAGTTCACCTTGCACCAATAGAGAGGCCCTCTGATGTTGATTCAGCTCC 1709  
Db 31295 CATGACCCCTCAGTTCACCTTGCACCAATAGAGAGGCCCTCTGATGTTGATTCAGCTCC 31354  
QY 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743  
Db 31355 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 31388

## RESULT 13

US-10-724-806-19  
; Sequence 19, Application US/10724806  
; Publication No. US20040248838A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELEY, RANDY D.  
; APPLICANT: APPARUNDARAM, SUBRAMANIAM  
; APPLICANT: PERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/10/724,806  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US/09/911,077A  
; PRIOR FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 119040  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: (2347)..(90873)  
; OTHER INFORMATION: N = A, C, G or T/U  
US-10-724-806-19

Query Match 28.8%; Score 502.8; DB 9; Length 119040;  
Best Local Similarity 87.1%; Pred. No. 3.5e-148;  
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AATGATCAGACAGAAATTTGTGGTCATGAGATCATCTGCTTGTTCGAGC 1169  
Db 30755 ACAGGCTTGGCAAGAAATCGTTGGTTATGCCAATCAGATGTTGTGTTGGAGC 30814  
QY 1170 ATTGCAACAGCATGGCTTGTGTCAGCAAGACTGTGTATGGGCTCTGTGAACCTGAGCTC 1229  
Db 30815 ATTGCAACAGCATGGCTTGTGTCAGCAAGACTGTGTATGGGCTCTGTGAACCTGAGCTC 30874  
QY 1230 TGACCTTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTTCACTCAAGAAAC 1289  
Db 30875 TGACCTTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTTCACTCAAGAAAC 30934  
QY 1290 CAACACTTATGGGAGGTTGCTGTTATATTTTGAACATCTCTGAAATTACTGAGG 1349  
Db 30935 CAACACTTATGGGAGGTTGCTGTTATATTTTGAACATCTCTGAAATTACTGAGG 30994  
QY 1350 AGAGCCATATCTTACTTGTGAGCCCTTAATCTTCAACCTGTTATTAATCTGACAAGAA 1409  
Db 30995 GAGGCCATATCTTACTTGTGAGCCCTTAATCTTCAACCTGTTATTAATCTGACAAGAA 31054

QY 1410 TGGTATATATATACAGAGGTTCCCATTTAAACCTCTCCATGGTTACCTCATTTCTTAAC 1469  
Db 31055 TGGTATATATATACAGAAATTTCCATTAAACACTTGCATGGTTACATCATTTCTTAAC 31114  
QY 1470 CAACATTGTGTTCTTATCTAGCAAGATATCTATTGGAAGGAACTTGGCTCCAA 1529  
Db 31115 CAACATTGTGATCTCTTATCTAGCAAGATATCTATTGGAAGGAACTTGGCTCCAA 31174  
QY 1530 ATTAGATGATTTGATGCTGTGTGCGAAGGCAAGTGAAGAGAACATGACAAGCCAT 1589  
Db 31175 ATTAGATGATTTGATGCTGTGTGCGAAGGCAAGTGAAGAGAACATGACAAGCCAT 31234  
QY 1590 TCTAGTCAGAAATGAAATATCAAAATTAATGAACCTTGCACCTGTGAACCTGGCAGAG 1649  
Db 31235 TCTGTCAAAAAGAAATTAATTAATGATGAACCTTGCACCTGTGAAGCCAGACAGAG 31294  
QY 1650 CCTAACCCCTCAGTTCACCTTGCACCAATAGAGAGGCCCTCTGATGTTGATTCAGCTCC 1709  
Db 31295 CATGACCCCTCAGTTCACCTTGCACCAATAGAGAGGCCCTCTGATGTTGATTCAGCTCC 31354  
QY 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743  
Db 31355 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 31388

## RESULT 14

US-09-911-077A-14  
; Sequence 14, Application US/09911077A  
; Publication No. US20030114399A1  
; GENERAL INFORMATION:  
; APPLICANT: BLAKELEY, RANDY D.  
; APPLICANT: APPARUNDARAM, SUBRAMANIAM  
; APPLICANT: PERGUSON, SHAWN  
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER cDNA  
; FILE REFERENCE: VBLT:008US  
; CURRENT APPLICATION NUMBER: US/09/911,077A  
; CURRENT FILING DATE: 2001-07-23  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 142299  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
; NAME/KEY: modified base  
; LOCATION: (1305)..(127835)  
; OTHER INFORMATION: N = A, C, G or T/U  
US-09-911-077A-14

Query Match 28.8%; Score 502.8; DB 3; Length 142299;  
Best Local Similarity 87.1%; Pred. No. 4e-148;  
Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1110 AATGATCAGACAGAAATTTGTGGTCATGAGATCATCTGCTTGTTCGAGC 1169  
Db 94673 ACAGGCTTGGCAAGAAATCGTTGGTTATGCCAATCAGATGTTGTGTTGGAGC 94732  
QY 1170 ATTGCAACAGCATGGCTTGTGTCAGCAAGACTGTGTATGGGCTCTGTGAACCTGAGCTC 1229  
Db 94733 ATTGCAACAGCATGGCTTGTGTCAGCAAGACTGTGTATGGGCTCTGTGAACCTGAGCTC 94792  
QY 1230 TGACCTTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTTCACTCAAGAAAC 1289  
Db 94793 TGACCTTGTCTACATCATCTTCCACAGCTGCTGTGTACTCTTCACTCAAGAAAC 94852  
QY 1290 CAACACTTATGGGAGGTTGCTGTTATATTTTGAACATCTCTGAAATTACTGAGG 1349  
Db 94853 CAACACTTATGGGAGGTTGCTGTTATATTTTGAACATCTCTGAAATTACTGAGG 94912  
QY 1350 AGAGCCATATCTTACTTGTGAGCCCTTAATCTTCAACCTGTTATTAATCTGACAAGAA 1409

```
Db 94913 GGAGCCATATCTGTATCTTACAGCCCTTGATCTTCTTACCCCTGCTATTAACCGATGATTA 94972
Qy 1410 TGGTATATATCAATCAGAGGTTCCCATTTAAACCTCCTCATGTTACCTCATTTCTTAC 1469
Db 94973 TGGTATATATATCAAGAAATTTCCATTTAAACACTGTCATGTTACATCTTCTTAAAC 95032
Qy 1470 CAACATTTGTGTTCTTATCTTACGCAAGTATCTATTTGAAAGTGAACCTTGCCTTCAAA 1529
Db 95033 CAACATTTGCACTCTCTATCTTACGCAAGTATCTATTTGAAAGTGAACCTTGCCTTCAAA 95092
Qy 1530 ATTATATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGSAACAAGCCAT 1589
Db 95093 ATTATATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGSAATPAAGCAAT 95152
Qy 1590 TCTAGTCAGAAATGAAATATCAATTAATGAATGAATCACTTGCACTTGCAAACTCGCAGAG 1649
Db 95153 TCTTGTCAAAAATGAAATATTAATTAATGATGATCACTTGCACTTGCAAACTCGCAGAG 95212
Qy 1650 CCTAACCTCAGTTCACTTTCAACCAATPAAGAGGCCCTCCTTGATGTTGATTCAGTCC 1709
Db 95213 CATGACCTCAGCTCAACTTTCAACCAATPAAGAGGCCCTCCTTGATGTTGATTCAGTCC 95272
Qy 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743
Db 95273 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 95306
```

## RESULT 15

```
US-10-724-806-14
; Sequence 14, Application US/10724806
; Publication No. US20040248838A1
; GENERAL INFORMATION:
; APPLICANT: BLAKELY, RANDY D.
; APPLICANT: APPARUNDARAM, SUBRAMANIAM
; APPLICANT: FERGUSON, SHAWN
; TITLE OF INVENTION: HUMAN AND MOUSE CHOLINE TRANSPORTER CDNA
; FILE REFERENCE: VBLT:008US
; CURRENT APPLICATION NUMBER: US/10/724,806
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US/09/911,077A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 142299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; NAME/KEY: modified base
; LOCATION: (1305)..(127835)
; OTHER INFORMATION: N = A, C, G or T/U
US-10-724-806-14
```

Query Match 28.8%; Score 502.8; DB 9; Length 142299;

Best Local Similarity 87.1%; Pred. No. 4e-148;

Matches 552; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy 1110 AAATGATCAGACAGAAATTTGTGGTCAATGAGATCACTGCTGTTGTTGGAGC 1169
Db 94673 ACAGGCTTCGACAAAGAAATGTTGGTTATGCAATCAAGTGTGTTTGGAGC 94732
Qy 1170 ATCTGCAACAGCCATGCTTGTCTGACGAAGACTGATGAGGCTTGATACCTGAGCTC 1229
Db 94733 ATCTGCAACAGCCATGCTTGTCTGACGAAGACTGATGAGGCTTGATACCTGAGCTC 94792
Qy 1230 TGACCTTGTCTAATCATCATTTTCCACAGCTGCTGTGTAATCTTTCATCAAGGAAC 1289
Db 94793 TGACCTTGTCTAATCATCATTTTCCACAGCTGCTGTGTAATCTTTCATCAAGGAAC 94852
```

```
Qy 1230 CAACATTAAGGGGAGTGTGTTATATTTTGGACTATTTCTGAAATTAATCGAAG 1349
Db 94853 CAACACTATGGGGCGTGGGAGGTTATGTTTCTGGCTCTTCTGAGAAATACGAGG 94912
Qy 1350 AGAGCCATATCTATCTTGCGAGCCCTTAATCTTCAACCTGGTTATTAATCTGACAAAGA 1409
Db 94913 GGAGCCATATCTGATCTTCAAGCCCTTGATCTTCAACCTGGCTATTAACCTGATGATTA 94972
Qy 1410 TGGTATATATCAATCAGAGTTCCCATTTAAACCTCCTCATGTTTACCTCATTTCTTAC 1469
Db 94973 TGGTATATATATCAAGAAATTTCCATTTAAACCTTGCATGTTATCATATCTTTAAC 95032
Qy 1470 CAACATTTGTGTTCTTATCTTACGCAAGTATCTATTTGAAAGTGAACCTTGCCTTCAAA 1529
Db 95033 CAACATTTGCACTCTCTATCTTACGCAAGTATCTATTTGAAAGTGAACCTTGCCTTCAAA 95092
Qy 1530 ATTATATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGSAACAAGCCAT 1589
Db 95093 ATTATATGATTTGATGCTGTTGTGCGAAGGACAGTGAAGAAACATGSAATPAAGCAAT 95152
Qy 1590 TCTAGTCAGAAATGAAATATCAATTAATGAATGAATGAATCACTTGCACTTGCAAACTCGCAGAG 1649
Db 95153 TCTTGTCAAAAATGAAATATTAATTAATGATGATCACTTGCACTTGCAAACTCGCAGAG 95212
Qy 1650 CCTAACCTCAGTTCACTTTCAACCAATPAAGAGGCCCTCCTTGATGTTGATTCAGTCC 1709
Db 95213 CATGACCTCAGCTCAACTTTCAACCAATPAAGAGGCCCTCCTTGATGTTGATTCAGTCC 95272
Qy 1710 GGAGGGGTCTGGGACTGAAGATTAATTACATGA 1743
Db 95273 AGAAGGGTCTGGGACTGAAGATTAATTACATGA 95306
```

Search completed: July 10, 2006, 14:50:57  
Job time : 2000 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:47 ; Search time 258 Seconds  
(without alignments)  
8783.456 Million cell updates/sec

Title: US-10-724-806-3

Perfect score: 1743

Sequence: 1 atgccttcacatgctgaag.....ctgaagataattacaatga 1743

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37.6	2.2	933	US-11-266-748A-359132	Sequence 359132,
2	37.6	2.2	933	US-11-266-748A-386971	Sequence 386971,
3	37.6	2.2	933	US-11-266-748A-442511	Sequence 442511,
4	37.6	2.2	1705	US-10-449-902-9396	Sequence 9396, AP
5	37	2.1	518	US-11-266-748A-14121	Sequence 14121, A
6	37	2.1	628	US-11-266-748A-75909	Sequence 75909, A
7	37	2.1	628	US-11-266-748A-109134	Sequence 109134,
8	37	2.1	628	US-11-266-748A-128720	Sequence 128720,
9	37	2.1	650	US-11-266-748A-48029	Sequence 48029, A
10	37	2.1	801	US-11-266-748A-10950	Sequence 10950, A
11	37	2.1	874	US-11-266-748A-177977	Sequence 177977,
12	37	2.1	902	US-11-266-748A-114600	Sequence 114600,
13	37	2.1	902	US-11-266-748A-130654	Sequence 130654,
14	37	2.1	1000	US-11-266-748A-119114	Sequence 119114,
15	37	2.1	1000	US-11-266-748A-161278	Sequence 161278,
16	37	2.1	1124	US-11-266-748A-178110	Sequence 178110,
17	37	2.1	1464	US-11-266-748A-227110	Sequence 227110,
18	37	2.1	1681	US-11-266-748A-84832	Sequence 84832, A
19	37	2.1	1681	US-11-266-748A-111914	Sequence 111914,
20	37	2.1	1681	US-11-266-748A-137643	Sequence 137643,
21	37	2.1	3214	US-11-266-748A-278774	Sequence 278774, A
22	36.6	2.1	1261	US-11-266-545-342	Sequence 342, App
23	36.6	2.1	165280	US-11-021-837-47	Sequence 47, App1
24	36	2.1	1768	US-11-266-748A-73373	Sequence 73373, A

25	36	2.1	1768	US-11-266-748A-107799	Sequence 107799,
26	36	2.1	1768	US-11-266-748A-126184	Sequence 126184,
27	35.8	2.1	755217	US-11-266-748A-39045	Sequence 29045, A
28	35.4	2.0	1791	US-10-449-902-4799	Sequence 4799, AP
29	35.4	2.0	2549	US-10-449-902-16653	Sequence 16653, A
30	35	2.0	804	US-11-266-748A-86191	Sequence 86191, A
31	35	2.0	804	US-11-266-748A-139002	Sequence 139002,
32	35	2.0	1000	US-11-266-748A-199625	Sequence 199625,
33	35	2.0	42790	US-11-266-748A-23980	Sequence 23980, A
34	34.8	2.0	10096	US-11-266-748A-59715	Sequence 59715, A
35	34.2	2.0	1099	US-11-266-748A-16467	Sequence 16467,
36	34	2.0	1000	US-11-266-748A-199378	Sequence 199378,
37	34	2.0	1962	US-11-266-748A-32112	Sequence 32112, A
38	34	2.0	2207	US-11-266-748A-25770	Sequence 25770, A
39	34	2.0	2876	US-10-505-928-468	Sequence 468, App
40	34	2.0	2876	US-11-341-567-1	Sequence 1, App1
41	34	2.0	5262	US-11-217-529-76367	Sequence 76367, A
42	34	2.0	9360	US-11-266-748A-28672	Sequence 28672, A
43	34	2.0	14544	US-11-266-748A-23683	Sequence 23683, A
44	34	2.0	17509	US-11-266-748A-28800	Sequence 28800, A
45	33.8	1.9	397	US-11-266-748A-37469	Sequence 37469, A

## ALIGNMENTS

RESULT 1	US-11-266-748A-359132	Application US/11266748A
1	Sequence 359132,	Publication No. US20060134663A1
2	GENERAL INFORMATION:	
3	APPLICANT: Johnston, Paul	
4	APPLICANT: Mulligan, Karl	
5	TITLE OF INVENTION: Transcriptome Microarray Technology and	
6	FILE REFERENCE: 55815-0102 (319189)	
7	CURRENT APPLICATION NUMBER: US/11/266,748A	
8	CURRENT FILING DATE: 2005-11-03	
9	PRIOR APPLICATION NUMBER: EP 04105479.2	
10	PRIOR FILING DATE: 2004-11-03	
11	PRIOR APPLICATION NUMBER: EP 04105482.6	
12	PRIOR FILING DATE: 2004-11-03	
13	PRIOR APPLICATION NUMBER: EP 04105483.4	
14	PRIOR FILING DATE: 2004-11-03	
15	PRIOR APPLICATION NUMBER: EP 04105507.0	
16	PRIOR FILING DATE: 2004-11-03	
17	PRIOR APPLICATION NUMBER: EP 04105485.9	
18	PRIOR FILING DATE: 2004-11-03	
19	PRIOR APPLICATION NUMBER: EP 04105484.2	
20	PRIOR FILING DATE: 2004-11-03	
21	PRIOR APPLICATION NUMBER: US 60/662,276	
22	PRIOR FILING DATE: 2005-03-14	
23	PRIOR APPLICATION NUMBER: US 60/700,293	
24	PRIOR FILING DATE: 2005-07-18	
25	NUMBER OF SEQ ID NOS: 483996	
26	SOFTWARE: PatentIn version 3.3	
27	SEQ ID NO 359132	
28	LENGTH: 933	
29	TYPE: DNA	
30	ORGANISM: Homo Sapiens	
31	US-11-266-748A-359132	
32	Query Match	2.2%; Score 37.6; DB 8; Length 933;
33	Best Local Similarity	45.1%; Pred. No 0.4; Indels 6; Gaps 2;
34	Matches 222; Conservative 0; Mismatches 264;	
35	119 AACCCATCATAGTGGGGCCCGTACATGTTGTTGTTGTTGTTTACCATGACAG 178	
36	279 AAGGACTACTTCCTGCGTGGAGGGGACATGTTGTTGTTGTTGTTGTTT 338	
37	179 CCACCTGGTTGAGGAGGCTACATGATGAGGACAGAGAGTGTATGGCCAGGTT 238	

Db 339 CCAGCAATGTTGGAGTGAATGACATTTTCATGTCCTGGCAGGGGTGACAGTGTCTGTACGGGCA 398  
Qy 239 GTGGCTAGCTGGGCTGATGACCCATTTGATATTTCTGTAGTCTAATTTTAGGTGTC 298  
Db 339 TTTCTGTA---TCACCTATGAACTTAATGCTTTGTTTCTGTGCTGATGTTGGCTGGA 455  
Qy 239 TGTTTTTCGCAAACTTATGCTTCCAAAGGATATGTGACTATGTTAGCCATTCACAA 358  
Db 456 TCTTCTTACCCATCTACATGTGCTGTGTCAGGTCCACAGATGCCAGAAATCTACGGAAGC 515  
Qy 359 AGATCTATGAAAGGCAATGGGTGGGCTGCTTTCATCCCTGCACATGAGGAGAGATG 418  
Db 516 GCTTGGTGGCATGCAATCCCATCATCTGCTGCTGATCTTACTATTTATCTACATCT 575  
Qy 419 T---CTGGGCTGACCAATTTTCTCTGCAATTAGGGGCCACCATCAGCGTATCATTTAG 475  
Db 576 TCACCAAGATCTCGGTAGACATGTATGACAGGTGCCATCTTACCCAGAGCTTTGGAC 635  
Qy 476 TGGATGTGAACATATCGGTCATGTCTCTGCACTCATTTGCCATTTCTTATACCTTAG 535  
Db 636 TGGATCTGTACTGTGCAATGTTGGGCTACTGGCCATCATCTGCTATATACAGGTTGCTG 695  
Qy 536 GTGGGCTCTACTCTGTGGCATATGATGTTGTCAGCTATTTCTGATTTTATAGAGC 595  
Db 696 GTGGGCTGGCTGTGTGATCTACAGGATGCCCTGACAGCGCTGATCATGCTTATAGAG 755  
Qy 596 TGTGATCAGTG 607  
Db 756 CGCTCACCCTTG 767

## RESULT 2

US-11-266-748A-386971  
; Sequence 386971, Application US/11266748A  
; Publication No. US2006013463A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcription Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 386971  
; LENGTH: 933  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-386971

Query Match 2.2%; Score 37.6; DB 8; Length 933;  
Best Local Similarity 45.1%; Pred. No. 0.4;  
Matches 222; Conservative 0; Mismatches 264; Indels 6; Gaps 2;  
Qy 119 AAGCATCATAGTCGGGGCGCTGACATTTGTTGTGTGTGTGTTTACATGACAG 178

Db 279 AAGGCTACTTCTGCTGAGGGGACATGCTGTGTGTCAGATGGTGCATCTCTGTTTG 338  
Qy 179 CCACCTGGGTTGGAGAGGCTACATCATATGGAGACAGAAAGCAGTGTATGGCCAGGTT 238  
Db 339 CCAGCAATGTTGGAGTGAATGACATTTTATGCTTGGCTGACAGGCTGTCTGTACGGGCA 398  
Qy 239 GTGGCTAGCTGGGCTGATGACCCATTTGATATTTCTGTAGTCTAATTTTAGGTGTC 298  
Db 456 TCTTCTTACCCATCTACATGTGCTGTGTCAGGTCCACAGATGCCAGAAATCTACGGAAGC 515  
Qy 299 TGTTTTTCGCAAACTTATGCTTCCAAAGGATATGTGACTATGTATACCATTCACAA 358  
Db 516 GCTTGGTGGCATGCAATCCCATCATCTGCTGCTGATCTTACTATTTATCTACATCT 575  
Qy 419 T---CTGGGCTGACCAATTTTCTCTGCAATTAGGGGCCACCATCAGCGTATCATTTAG 475  
Db 576 TCACCAAGATCTCGGTAGACATGTATGACAGGTGCCATCTTATCCAGAGCTTTGGAC 635  
Qy 476 TGGATGTGAACATATCGGTCATGTCTCTGCACTCATTTGCCATTTCTTATACCTTAG 535  
Db 636 TGGATCTGTACTGTGCAATGTTGGGCTACTGGCCATCATCTGCTATATACAGGTTGCTG 695  
Qy 536 GTGGGCTCTACTCTGTGGCATATGATGTTGTCAGCTATTTCTGATTTTATAGAGC 595  
Db 696 GTGGGCTGGCTGTGTGATCTACAGGATGCCCTGACAGCGCTGATCATGCTTATAGAG 755  
Qy 596 TGTGATCAGTG 607  
Db 756 CGCTCACCCTTG 767

## RESULT 3

US-11-266-748A-442511/c  
; Sequence 442511, Application US/11266748A  
; Publication No. US2006013463A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcription Microarray Technology and  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 442511  
; LENGTH: 933  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-442511

Query Match 2.2%; Score 37.6; DB 8; Length 933;  
Best Local Similarity 45.1%; Pred. No. 0.4;  
Matches 222; Conservative 0; Mismatches 264; Indels 6; Gaps 2;

```
OY 119 AAGCATCATAGTCGGGGGCGTGAACATGTTGTTGGTGGTGGTTTACCATGACAG 178
DB 655 AAGGCTACTTCTGGCTGGAGGGGACATGTGTGGGCGAGTGGGTGATCTTGTGG 596
OY 179 CCACCTGGGTGGAGAGGCTCATCAATGAGGACAGCAAGCATGTAATGGGCAAGTT 238
DB 595 CCAGCAATGTGGAGATGACATTTTATGCGCTGGCAGGGTCAAGGTGCTGTACAGGGA 536
OY 239 GTGTCTAGCTTGGGCTCATGACCATTTGATTAATCTGAGTCTAATTTTAAAGTGTG 298
DB 535 TTTCTGTA---TCAGTTATGAACCTTAATGAGCTTTTCTGCTGATGTTGGCTGGA 479
OY 299 TGTTTTGGGAACCTTATGCGTTCCAGGATATGTGACTATGTTTGAACCATTCACAC 358
DB 478 TCTTCTTACCATCTACATTTGCTGGTCAAGTCAACGATGCGAATACCTACGGAAGC 419
OY 359 AGATCTATGGAAGGCAATGGGTGGGCTGCTTATCCCTGCACTGATGGAGAGATGT 418
DB 418 GCTTGGTGGCATCAGAAATCCCATCATCTGCTGCTACTTACCTATTATCTACATCT 359
OY 419 T---CTGGGCTGAGCAATTTTCTCTGCAATTAAGGGGCAACATCAGCGTATCATTTGATG 475
DB 358 TCACCAAGATCTCGTGAACATGTATGACAGTGCATCTTTCACGACGATCTTGGACAC 299
OY 476 TGGATGTGAACATATGCTGATTTGTCTGCACTCATTTGCTTATTAACCTTAATGG 535
DB 298 TGGATCTGACCTGGCCATAGTTGGGCTACTGGCCATCACTGCTGATATACAGGTTGCTG 239
OY 536 GTGGGCTCTACTCTGTGGGATATGATGTTGTCCAGCTATTCGATTTTATATGAGC 595
DB 238 GTGGGCTGGCTGTGATCTACAGGATGCCCTGAGCGCTGATCATGCTTATAGAGG 179
OY 596 TGTGATCAGTG 607
DB 178 CGCTCACCTTTG 167
```

## RESULT 4

US-10-449-902-9396  
; Sequence 9396, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agrobiological Sciences.  
; APPLICANT: Bio-oriented Technology Research Advancement Institution.  
; APPLICANT: The Institute of Physical and Chemical Research.  
; APPLICANT: Foundation for Advancement of International Science.  
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
; FILE REFERENCE: MOA-A0205Y1-US  
; CURRENT APPLICATION NUMBER: US/10/449, 902  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: JP 2002-203269  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: JP 2002-383870  
; PRIOR FILING DATE: 2002-12-11  
; NUMBER OF SEQ ID NOS: 56791  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9396  
; LENGTH: 1705  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AK106744  
; DATABASE ENTRY DATE: 2002-08-28  
US-10-449-902-9396

Query Match 2.2%; Score 37.6; DB 6; Length 1705;  
Best Local Similarity 57.8%; Pred. No. 0.58;  
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
OY 1276 TTCATCAAGGACCACTTATGGGCGTGGTGTATATTTGGACTATTCCTG 1335
DB 413 TTCATTCAGCAAAATGCTGTGGTGCAGATCTGTGATCTTTGGACTTAAAG 472
OY 1336 AGAATTACTGAGAGGACCATATCTATCTTGCAGCCCTTATCTTACCTCG 1391
DB 473 GCATTGATTTAGCGCCGCTTATCATATCTTCAAGTTACCATCATCATTTGG 528
```

## RESULT 5

US-11-266-748A-14121  
; Sequence 14121, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105482.6  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105483.4  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105507.0  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105485.9  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: EP 04105484.2  
; PRIOR FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US 60/662,276  
; PRIOR FILING DATE: 2005-03-14  
; PRIOR APPLICATION NUMBER: US 60/700,293  
; PRIOR FILING DATE: 2005-07-18  
; NUMBER OF SEQ ID NOS: 483996  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 14121  
; LENGTH: 518  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-11-266-748A-14121

Query Match 2.1%; Score 37; DB 8; Length 518;  
Best Local Similarity 64.7%; Pred. No. 0.42;  
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
OY 185 GGGTTGGAGGAGGACATCATATGGAGCAGAGAGAGTGTATGGGCCAGGTTGTGTC 244
DB 371 GAGTTGGCCGAGAGATGATGATGTGGCAGAGAAAGCCGTTGCAAGCCCAAGTTGCTC 430
OY 245 TAGCTTGGGCTCATGACCCCATTTGG 269
DB 431 TAGGACAGGACCTTTGGCCCTTTG 455
```

## RESULT 6

US-11-266-748A-75909  
; Sequence 75909, Application US/11266748A  
; Publication No. US20060134663A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkin, Paul  
; APPLICANT: Johnston, Patrick  
; APPLICANT: Mulligan, Karl  
; TITLE OF INVENTION: Transcriptome Microarray Technology and  
; TITLE OF INVENTION: Methods of Using the Same  
; FILE REFERENCE: 55815-0102 (319189)  
; CURRENT APPLICATION NUMBER: US/11/266,748A  
; CURRENT FILING DATE: 2005-11-03  
; PRIOR APPLICATION NUMBER: EP 04105479.2

```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 75909
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-75909
```

```
Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```
QY 185 GGGTTGAGAGAGGCTACATCATGATGGACAGACAGTGTATGGCCAGGTTGTGTC 244
DB 484 GAGTTGGCGGAGAGATGATGTGGCGACGAGAAAGCCCTTGACGCCACCTGTGTGCTC 543
QY 245 TAGCTTGGGCTCATGCACCCATTGG 269
DB 544 TAGCGAGGAGCCTTGGCCCTTTG 568
```

```

RESULT 7
US-11-266-748A-109134
; Sequence 109134, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 109134
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-109134
```

```

Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 185 GGGTTGAGAGAGGCTACATCATGATGGACAGACAGTGTATGGCCAGGTTGTGTC 244
DB 484 GAGTTGGCGGAGAGATGATGTGGCGACGAGAAAGCCCTTGACGCCACCTGTGTGCTC 543
QY 245 TAGCTTGGGCTCATGCACCCATTGG 269
DB 544 TAGCGAGGAGCCTTGGCCCTTTG 568
```

```

RESULT 8
US-11-266-748A-128720/c
; Sequence 128720, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 128720
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-128720
```

```

Query Match      2.1%; Score 37; DB 8; Length 628;
Best Local Similarity 64.7%; Pred. No. 0.48;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 185 GGGTTGAGAGAGGCTACATCATGATGGACAGACAGTGTATGGCCAGGTTGTGTC 244
DB 145 GAGTTGGCGGAGAGATGATGTGGCGACGAGAAAGCCCTTGACGCCACCTGTGTGCTC 86
QY 245 TAGCTTGGGCTCATGCACCCATTGG 269
DB 85 TAGCGAGGAGCCTTGGCCCTTTG 61
```

```

RESULT 9
US-11-266-748A-48029
; Sequence 48029, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
```





```
/ Sequence 114600, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 114600
/ LENGTH: 902
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-114600

Query Match      2.1%; Score 37; DB 8; Length 902;
Best Local Similarity 64.7%; Pred. No. 0.6;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1177 ACAGCATTGCTTCTTCAGTGAAGACTGTATGGGCTCTGTACCTGAGCTTGACCTT 1236
    |||||
DB 523 AGAGCTTCGCTTCCAGTCGAGGATTATTAGCGGCTCTTCTTCTTACGCTTGAGCTTC 464
    |||||

OY 1237 GCTACATCATCATCTTCCACAGC 1261
    |||||
DB 463 CATTTCCGATCATGATTCATCCAGC 439
    |||||

RESULT 13
US-11-266-748A-120654
/ Sequence 120654, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
```

```
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 120654
/ LENGTH: 902
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-120654

Query Match      2.1%; Score 37; DB 8; Length 902;
Best Local Similarity 64.7%; Pred. No. 0.6;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 1177 ACAGCATTGCTTCTTCAGTGAAGACTGTATGGGCTCTGTACCTGAGCTTGACCTT 1236
    |||||
DB 380 AGAGCTTCGCTTCCAGTCGAGGATTATTAGCGGCTCTTCTTACGCTTGAGCTTC 439
    |||||

OY 1237 GCTACATCATCATCTTCCACAGC 1261
    |||||
DB 440 CATTTCCGATCATGATTCATCCAGC 464
    |||||

RESULT 14
US-11-266-748A-119114
/ Sequence 119114, Application US/11266748A
/ Publication No. US2006013463A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: Patentin version 3.3
/ SEQ ID NO: 119114
/ LENGTH: 1000
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-11-266-748A-119114

Query Match      2.1%; Score 37; DB 8; Length 1000;
Best Local Similarity 64.7%; Pred. No. 0.64;
Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY 185 GGGTGGAGGAGGCTACATCATGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 244
    |||||
DB 856 GAGTTGGCCGAGAGAGATGATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
    |||||

OY 245 TAGCTTGGGCTTCATGACACCATTTGG 269
    |||||
DB 916 TAGGAGAGGAGACTTTGGCCCTTTG 940
    |||||
```

Tue Jul 11 13:42:50 2006

us-10-724-806-3.рnpbn

Page 7

## RESULT 15

US-11-266-748A-161278/c

Sequence 161278, Application US/11266748A  
Publication No. US20050134663A1

Publication No. US20060134663A1

;; GENERAL INFORMATION:  
; APPLICANT: Harkin,

APPLICANT: Johnston, Paul

APPLICANT: Mulligan, Karl

TITLE OF INVENTION: Transcriptome Microarray Technology and

```
; TITLE OF INVENTION: Methods of Using the Same  
FILE NUMBER: 6017-0100 (3-10-80)
```

FILE REFERENCE: 55815-0102 (319189)  
CURRENT ADDICTION NUMBER: HSC/11/3266 7482

CURRENT APPLICATION NUMBER: US/11/266,748A  
CURRENT FILING DATE: 2005-11-03

PRIOR APPLICATION NUMBER: EP 04105479.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105482.6

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105483.4

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105507.0  
PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105485 9

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: EP 04105484.2

PRIOR FILING DATE: 2004-11-03

PRIOR APPLICATION NUMBER: US 60/662,276

PRIOR FILING DATE: 2005-03-14

PRIOR APPLICATION NUMBER: US 60/700,293  
PRIOR FILING DATE: 2005-07-18

PRIOR FILING DATE: 2005-07-18  
NUMBER OF SEQ ID NOS: 483996

```

; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3

```

```

; SOFTWARE: FACEMILL VERSION 3.3
; SEQ ID NO 161278

```

```

;      LENGTH: 1000
;

```

TYPE: DNA

ORGANISM: Homo Sapiens

US-11-266-748A-161278

Query Match 2 1%: Score 37

Query Match	2.1%;	Score 37
Best Local Similarity	64.78;	Pred. Nccl

Best Local Similarity	64.7%	Pred. NCC
Matches	55;	Conservative
		0;
		Mismatches

1. The first group of people who are not allowed to enter the country are those who are considered to be a threat to national security. This includes anyone who is involved in espionage, sabotage, or other activities that could harm the country's interests.

185 GGGTTGGAGGAGGCTACATCAATGGGAC

\_\_\_\_\_

Db 145 GAGTTGCCCGAGGATGGATGATTGTGGC

[illegible]

245 TAGCTGGGCTCATGCACCCATTGG 26

Search completed: July 10, 2006, 23:23:51  
Job time : 261 secs

**This Page Blank (uspio)**

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocooperation Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 10, 2006, 14:01:37 ; Search time 350 Seconds  
(without alignments)  
9318.122 Million cell updates/sec

Title: US-10-724-806-3

Perfect score: 1743  
Sequence: 1 atgccttcacatggaag.....ctgaagaattacatga 1743

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued Patents NA,\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6 COMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PC/US COMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PC COMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfill1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1378.2	79.1	US-09-657-252-1	Sequence 1, Appl
2	237.6	13.6	US-09-974-300-501	Sequence 501, App
3	138.8	8.0	US-09-270-767-1720	Sequence 1720, Ap
4	138.8	8.0	US-09-270-767-17002	Sequence 17002, A
5	45.8	2.6	US-08-232-463-14	Sequence 14, Appl
6	40	2.3	US-09-543-681A-822	Sequence 822, Appl
7	38.8	2.2	US-09-662-254B-26	Sequence 26, Appl
8	38.2	2.2	US-09-328-352-2245	Sequence 2245, Ap
9	37.8	2.2	US-09-640-198D-3	Sequence 3, Appl
10	37.8	2.2	US-08-595-553A-1	Sequence 1, Appl
11	37.8	2.2	US-09-995-007-1	Sequence 1, Appl
12	37.8	2.2	US-09-949-016-15637	Sequence 15637, A
13	37	2.1	US-09-513-999C-2240	Sequence 2240, Ap
14	37	2.1	US-09-539-333D-1	Sequence 1, Appl
15	37	2.1	US-09-679-409-1	Sequence 1, Appl
16	36.6	2.1	US-09-949-016-204222	Sequence 204222,
17	36.6	2.1	US-09-949-016-204223	Sequence 204223,
18	36.6	2.1	US-09-949-016-17511	Sequence 17511, A
19	36	2.1	US-09-774-639-63	Sequence 63, Appl
20	35.8	2.1	US-09-949-451-932	Sequence 932, App
21	35.8	2.1	US-09-949-016-17594	Sequence 17594, A
22	35.8	2.1	US-09-949-016-17417	Sequence 17417, A
23	35.8	2.1	US-09-949-016-12249	Sequence 12249, A

24	35.6	2.0	227	3	US-09-513-999C-9321	Sequence 9321, Ap
25	35.4	2.0	1689	3	US-09-108-010B-14	Sequence 14, Appl
26	35.4	2.0	1689	3	US-09-758-652-14	Sequence 14, Appl
27	35.4	2.0	1689	3	US-10-684-651-14	Sequence 14, Appl
28	35.4	2.0	58597	3	US-09-949-016-13598	Sequence 13598, A
29	35.4	2.0	139562	3	US-09-949-016-13451	Sequence 13451, A
30	35.2	2.0	23606	3	US-09-949-016-16398	Sequence 16398, A
31	35	2.0	821	3	US-09-786-240-30	Sequence 30, Appl
32	35	2.0	1048	3	US-09-533-559-6603	Sequence 6603, Ap
33	35	2.0	2613	3	US-09-976-594-1015	Sequence 1015, Ap
34	34.8	2.0	832	3	US-09-621-976-2813	Sequence 2813, Ap
35	34.6	2.0	601	3	US-09-949-016-88583	Sequence 88583, A
36	34.6	2.0	88268	3	US-09-949-016-14178	Sequence 14178, A
37	34.6	2.0	174259	3	US-09-949-016-11968	Sequence 11968, A
38	34.6	2.0	174262	3	US-09-949-016-14259	Sequence 14259, A
39	34.4	2.0	601	3	US-09-949-016-148743	Sequence 148743, A
40	34.4	2.0	1141	3	US-09-806-708B-22	Sequence 22, Appl
41	34.4	2.0	2028	3	US-10-162-012-28	Sequence 28, Appl
42	34.4	2.0	2034	3	US-09-949-016-1691	Sequence 1691, Ap
43	34.4	2.0	2326	3	US-10-162-012-26	Sequence 26, Appl
44	34.4	2.0	24593	3	US-09-949-016-13433	Sequence 13433, A
45	34.2	2.0	300	3	US-09-543-681A-1278	Sequence 1278, Ap

#### ALIGNMENTS

102(e)

RESULT 1		US-09-657-252-1	
Sequence 1, Application US/09657252			
Patent No. 6500643			
GENERAL INFORMATION:			
APPLICANT:	Wu, Dong-Hai		
APPLICANT:	Gu, Yunrong		
APPLICANT:	Millard, William		
APPLICANT:	He, Yun-Je		
TITLE OF INVENTION:	Human High Affinity Choline Transporter CDNA		
FILE REFERENCE:	MBH00-639		
CURRENT APPLICATION NUMBER:	US/09/657,252		
CURRENT FILING DATE:	2000-09-07		
NUMBER OF SEQ ID NOS:	6		
SOFTWARE:	PatentIn version 3.0		
SEQ ID NO 1			
LENGTH:	1743		
TYPE:	DNA		
ORGANISM:	Homo sapiens		
FEATURE:			
NAME/KEY:	CDS		
LOCATION:	(1)..(1743)		
US-09-657-252-1			
Query Match	79.1%;	Score 1378.2;	DB 3; Length 1743;
Best Local Similarity	86.9%;	Pred. No. 0;	
Matches 1515; Conservative	0;	Mismatches 228;	Indels 0; Gaps 0;
QY	1	ATGCTTTCATGGAAGAGCTGATATATCTCTTACCTCTTATATTTCTG	60
DB	1	ATGCTTTCATGGAAGAGCTGATATATCTCTTACCTCTTATATTTCTG	60
QY	61	GTGGAATATGGCTGATGAGAAACCAAAACAGCGCAACCCAGAGCGAGTAA	120
DB	61	GTGGAATATGGCTGATGAGAAACCAAAACAGCGCAACCCAGAGCGAGTAA	120
QY	121	GCATCATATGCTGGGGGCGGTGACATGTTGTTGTTGTTTACATGACAGCC	180
DB	121	GCATCATATGCTGGGGGCGGTGACATGTTGTTGTTTACATGACAGCT	180
QY	181	ACCTGGGTTGAGAGGCTATCATATGAGACAGCAAGAGTATGAGCCAGTTGT	240
DB	181	ACCTGGGTTGAGAGGCTATCATATGAGACAGCTGAGACAGTTATGTAACAGTTAT	240
QY	241	GGTGTACCTGGGCTCATGACCCCATTTGATATTTCTGAGTCTPATTTTAGGTGCTG	300

Db 241 GGCCTAGCTTGGGCTCAGGACCAATTGGATATCTCTTAGTCTGATTTTAAAGTGGCTG 300  
QY 301 TTTTGTGGAAAACCTATGCGTTCGAAGGATATGTACTATGTAGACCATTCAACAG 360  
Db 301 TTTCTTGGAAAACCTATGCGTTCGAAGGATATGTAGACATGTAGACCCGTTTACAGCA 360  
QY 361 ATCTATGGAAAAGCGATGGGTGGGCTGCTCTTCATCTCCCTGCACCTGATGGGAGATGTTG 420  
Db 361 ATCTATGGAAAAGCGATGGGCGGACCTCTGTTTATCTCCGACTGATGGGAGAAAATGTTTC 420  
QY 421 TGGGCTGAGCAATTTTCTCTGATTAAGGGGCCACCATCAGCGTATCATGATGAT 480  
Db 421 TGGGCTGAGCAATTTTCTCTGATTAAGGGGCCACCATCAGCGTATCATGATGAT 480  
QY 481 GTGAACATATCGGTGATGTTCTCTGACCTCATTTGCCATTTCTTATTAACCTAGTGGGTGG 540  
Db 481 ATGCACATTTCTGTATCATCTCTGCACATTCATGACATCTGTACACACTGTGTGGAGGG 540  
QY 541 CTCTACTCTGAGGCGATATCTGATGTTGTCAGCTATTTCTGATTTTATAGACTGTGG 600  
Db 541 CTCTACTCTGAGGCGATATCTGATGTTGTCAGCTATTTCTGATTTTATAGACTGTGG 600  
QY 601 ATCAGTGTCCCTTTTGCCCTGTACATCTCTGAGTCAACCGACATCGGATTCACAGCTGTG 660  
Db 601 ATCAGTGTCCCTTTTGCCCTGTACATCTCTGAGTCAACCGACATCGGATTCACAGCTGTG 660  
QY 661 CATGCTAAATACCAAGTCCCTGCTGGGAGACCATTTGAATCAGTTGAAGCTTACACCTGG 720  
Db 661 CATGCTAAATACCAAGCCGTGGCTGGGAGCTGTTGATCTCATCTGAAGTCTACTCTTGG 720  
QY 721 CTGTATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCACTTCCAGAG 780  
Db 721 CTGTATATTTTCTGTATATGATGCTGGGTGGAAATCCCATGGCAAGCCACTTCCAGAG 780  
QY 781 GTCTCTCTCTTCATCTCTCAGCCACTATGTCTCAGGTACTCTCTCTCTGAGCTTTTGGG 840  
Db 781 GTCTCTCTCTCTCTCTCAGCCACTATGTCTCAGGTACTCTCTCTCTGAGCTTTTGGG 840  
QY 841 TGCCTGTGATGGCTCTACCCGCCATATGATAGAGACTATTTGAGCTTCCACAGACTGG 900  
Db 841 TGCCTGTGATGGCTCTACCCGCCATATGATAGAGACTATTTGAGCTTCCACAGACTGG 900  
QY 901 AACCGACGCTTACGGGGATCCAGATCCCAAGACATTAAGAGAGAAAGCAATGATTC 960  
Db 901 AACCGACGCTTACGGGGATCCAGATCCCAAGACATTAAGAGAGAAAGCAATGATTTTA 960  
QY 961 CCGATGCTTGTGACATCTGCTGCTGTATCATCTCTCTTGGGGCTTGTGCTGTT 1020  
Db 961 CCGATGCTTGTGACATCTGCTGCTGTATCATCTCTCTTGGGGCTTGTGCTGAGTT 1020  
QY 1021 TCAGCTGCTGATGTTCTCTGAGCTGATCTGTCATCTCTGCGGCGAGTTCTATGTTTGGT 1080  
Db 1021 TCAGCTGCTGATGTTCTCTGAGCTGATCTGTCATCTCTGCGGCGAGTTCTATGTTTGGT 1080  
QY 1081 CGGAATATCTACCACTTTCCTTCAGACAAATGCAATCGAAGAGAAATTTGGGGCTC 1140  
Db 1081 CGGAATATCTACCACTTTCCTTCAGACAAATGCTTCGGAAGAAAGAAATTCGTTGGGTT 1140  
QY 1141 ATGAGATCATCTGTCTGTGTGTCGAGAGCATCTGCAACAGCCATGGCTTGTGTCAGAG 1200  
Db 1141 ATGAGATCATCTGTGTGTGTCGAGAGCATCTGCAACAGCCATGGCTTGTGTCAGAGAA 1200  
QY 1201 ACTGTGTATGGGCTGTGGTACTGAGCTTGAACCTTGTCTTACATCATCTTCCACAG 1260  
Db 1201 ACTGTGTATGGGCTGTGGTACTGAGCTTGAACCTTGTCTTACATCATCTTCCCCAG 1260  
QY 1261 CTGCTGTGTATCTCTTCAATCAAAAGAACCAACATTTATGGGGAGTTGCTGATATAT 1320  
Db 1261 CTGCTGTGTATCTCTTGTATAGGAAACCAACCTTATGGGGAGCTGAGGATATATGTT 1320  
QY 1321 TTTGGACTATTTCTTGAAGATTACTGGAGAGAGCAATCTTATCTTGGAGCCCTTAATC 1380  
Db 1321 TCTGGCTCTTCTTGAAGATTACTGGAGAGAGCAATCTTATCTTCAAGCCCTTGAATC 1380

QY 1381 TTTCAACCTGGTTATTTACTCTGACAAAGATGTATATACAATCAGAGGTCCCAATTAA 1440  
Db 1381 TTTCAACCTGGTTATTTACTCTGATGATATATATATATCAAAATTTCCATTTAA 1440  
QY 1441 ACTCTCTCCATGTTTACTCTCATCTCTTTCATTCGAACATTTGTGTTTCTTATCTAGCCAAAT 1500  
Db 1441 ACATCTGCAATGTTTACTCATCTCTTTCATTCGAACATTTGTGTTTCTTATCTAGCCAAAT 1500  
QY 1501 CTATTTGAAAGTGAACCTTTCCTCCAAATTTAGATGATTTGATGCTGTTGTCCAGAG 1560  
Db 1501 CTATTTGAAAGTGAACCTTTCCTCCAAATTTAGATGATTTGATGCTGTTGTCCAGAG 1560  
QY 1561 CACAGTGAAGAGACATGAGACCAATCTTACTGACAGAAATGAAATATCAATTAAT 1620  
Db 1561 CACAGTGAAGAGACATGAGATTAACCAATCTTGTCAAAATATGAAATATTAATTAAT 1620  
QY 1621 GAACCTTGACCTGTGAACCTTCGGAGAGCCATCCCTCAGTTCACTTCAACCAATAAG 1680  
Db 1621 GAACCTTGACCTGTGAACCTTCGGAGAGCCATCCCTCAGTTCACTTCAACCAATAAG 1680  
QY 1681 GAGGCTCTCTTATGATTTTCCAGTCCGAGAGGGGTCTGGGACTGAAGATTAATTACA 1740  
Db 1681 GAGGCTCTCTTATGATTTTCCAGTCCGAGAGGGGTCTGGGACTGAAGATTAATTACA 1740  
QY 1741 TGA 1743  
Db 1741 TGA 1743

RESULT 2  
US-09-974-300-501  
; Sequence 501, Application US/09974300  
; Patent No. 7018794  
; GENERAL INFORMATION:  
; APPLICANT: Berka, Randy M.  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 10085.500-US  
; CURRENT APPLICATION NUMBER: US/09/974,300  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 09/680,598  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/279,526  
; NUMBER OF SEQ ID NOS: 8481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 501  
; LENGTH: 1461  
; TYPE: DNA  
; ORGANISM: Bacillus licheniformis  
US-09-974-300-501

Query Match 13.6%; Score 237.6; DB 5; Length 1461;  
Best Local Similarity 51.9%; Pred. No. 3.3e-67;  
Matches 670; Conservative 0; Mismatches 574; Indels 48; Gaps 4;

QY 124 ATCATATGTCGGGGGCGCTGACATTTGTTGTTGGTGTGTTTATCCATGACAGCCACC 183  
Db 98 ATGATCGTGGCAAAAAGAGATGCGTTTTCGTGGAAATGATGATGATGCGCGCAACA 157  
QY 184 TGGGTGAGAGAGCTATCATCAATGGGACAGACAGAGATGATAGGCGCAGGTTGGT 243  
Db 158 TGGGTGAGAGAGCTATCATCAATGGGACAGACAGAGATGATAGGCGCAGGTTGGT 243  
QY 244 CTAGCTTGGGCGTACGACCCATGAGATATCTGAGCTTAATTTTATGTTGGTGTGTTT 303  
Db 212 CTGATCTGGGCGCAAGCCCTTGGGCTTACGATGATGAGCTGATATATGCGGATTTTCT 271  
QY 304 TTTGGAAACATATGCTTTCAGAGGATATGTGACTATGTAGACCCATTCAACAGATC 363  
Db 272 TTTGGCAAAAATATGCGCGTATCAATTTATGACATTTATGATCTCTGGAACAAGGC 331

QY	364	TATGAAAGCCCATGGGAGGCGCTGCTCTTCAATCCCTGCAGTCAGTGGAGAGATGTTCTGG	423
Db	332	TTGGGTAAAGCGGATGCGCGCGTTTATATATACCGGCGCTGTAGAGAAATGTTTGG	391
QY	424	GCTGACGCAATTTCTCTGCATTAAGGGCCACCATCAGCTGATCATTTGATGTGATGTG	483
Db	392	AGCCCGCGGATCTTAACGGCATTTGGGCGAGCTTTGGAAAGATCTCAAAATATCATTTTC	451
QY	484	AACATATGCGATGTGTTCTGCACACTCATGSCATTTTATATACCCTAGTGGGAGGCTC	543
Db	452	CAAAACGTGATTAATTTCTTGGCCGATGATCGCATCGTTATACGGTGGCCGAGGCAATG	511
QY	544	TACTCTGGCATTAATCATGATGTGTCCAGTCAATTTCTGCATTTTATAGACTGTGATC	603
Db	512	TGGCAGATGTCTTTCACAGATGTCTTTCAAAATGATGTGATTTTCTCGGCGTGTTTTAA	571
QY	604	AGTGTCCCTTTTGGCCCTGTACATCTGCACTACCGACATCGGATTCAGACAGTGTG---	660
Db	572	GTCGTCCCATTTGATGTGCAAAATGTCGATCTCTCGATACGATATGGGCAAAATTAACAG	631
QY	661	-----CATGCTAAATTAACAGATGCCGATGCCGTGGGAAACATTAAT	700
Db	632	CACGATTTCCGACAGCGCCAAATTTGCTTCCGCGCTTGGACGAGCTGGAAAAACCCGAT	691
QY	701	CAGTTGAA-----GTCTACACCTGGCTTGATTAATTTTCTGTTATTTGATGCTGGGTGAATC	756
Db	692	TGGGGAACCTGTTTTGGAACTGGGGGATTAATGCGTCTCTTTAATTTTCCGAGGATATC	751
QY	757	CCATGGCAGGCTTAATCTCCAGAGGTCCTCTCTTCAATCTCAGCCACTATGCTCAGSTA	816
Db	752	GCAATGGCAGGTATCTTTCACGCGCTTTCTTTCGGGAAATGGAAAAAGGCGCCCATGTGG	811
QY	817	CTGTCTCTTCCTGGGAGCTTTTGGGAGCTGGATGAGTGGCTCTACCCGCGCATATGATAGA	876
Db	812	CAGTTCGATTAATTTGCCGAGATGATCTGCATCAATGCCGCAATTCGCTGGGTATATCATG	871
QY	877	GCTATTTGAGCTTCCACAGACTGGAACCAAGACTGCTTACGGGATATCCAGATCCAAAGCT	936
Db	872	GCTGCGGAAACAGTACCGATTTGGAGCTGTTCGGAGGAGGCGCTCCGATATACCCGCG	931
QY	937	AAGAGGAGAGACATATATCTCCCATCGTTTCTGCAGTACTGTGCCCTGTGTATCATC	996
Db	932	-----ATGATTTTGGCCGAAACGCTTGGGTATTTTACCGCTCGCATCAAGCCGAAGGCC	976
QY	997	TCCTTCTTTGGGCTGTGCTGTTCACGCTGCTGATGCTCTCAGCTGACCTGCGCATC	1056
Db	977	GCAGGCTTCGGCTTGGGTGCAATCGCACAGCCGTCATGTACAAGATGGAATCATCATGAT	1038
QY	1057	CTGTGGCGGAGTTCTAATGTTTGTCTCGGAATCTTACAGCTTTCCTTCAGACAAATGCA	1118
Db	1037	CTATTCGCGATCATCAATGACCGGCAATGAAATTTTACCGCTCGCATCAAGCCGAAGGCC	1098
QY	1117	TCAGACAGGAAATGTGTGGGTACATGAGATCACTGTGCTTGTGTGGAGCATCTGCA	1176
Db	1097	ACCCAAAACAGCTGCAAAAAGTCTCTAAAGCTTAATCATTTTGTGTGGCGCGGAGCA	1158
QY	1177	ACAGCCATGAGCTTTGCTGACGAGAAGACTGTATATGAGGCTCTGGTACTGAGCTTGACCT	1238
Db	1157	GCGGTCAATCGCGTGAATGTCAAAAAGCTTTATACTTATATGATATTTGGCTTCGGATTTA	1218
QY	1237	GTCTACATCATCATCTTCCACAGCTGCTGTGTATCTTTCATCAAAAGAACCAACT	1298
Db	1217	GTATATTCATCTTTTTCGCCAGTTAAACAATGCGCTCTTTTATATAAAGAGCAATCTT	1278
QY	1297	TATGGGCGAGTGTGCTGTATATTTTGGACTATTTCTGAGAAATTAATCTGAGAGAGCCA	1358
Db	1277	TACGGGTGATGTGCTGATTTTGCAGTTGCAAGTCAATCTTGAGGCTCGGCGGTGTGAACC	1338
QY	1357	TATCATATCTTGCAGCCCTTAATCTTTACCC	1388
Db	1337	GCATTCGGCATTTCCGCGCTTCTGCGGATCC	1368

```

RESULT 3
US-09-270-767-1720
; Sequence 1720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7336-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1720
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1720

```

Query Match	8.0% ; Score 138.8 ; DB 3 ; Length 324 ;
Best Local Similarity	65.5% ; Pred. No. 5.1e-35 ;
Matches 203 ; Conservative 0 ; Mismatches 107 ; Indels 0 ; Gaps 0 ;	
QY	808 GCTCAGGACTGTCCTCTCTCGGACAGCTTTGGGTCCTGTGATGAGCTCTACCCGCCATA 867
Db	15 GCCCAGCTTCGTCTCTATGTTGACCGCGGAGATGATTTTGAATGGCCATTTCCCGGTTG 74
QY	868 TGCATAGAGCTATTGAGCTTCCACAGACTGGAACCAAGACTGCCCTACGGGTATTCAGAT 927
Db	75 CTCATCGGACGATATGGCCAAAGCTACACCTTTGAACAGACAGATTTCMAAGGACCCTAT 134
QY	928 CCCAAGACTAAGAGGAGAGCAGATATATCTCCGATCTGTCGAGTACTCTGCGCT 987
Db	135 CCCCTGACCGTGACGAGACGAGATATATCTGCCATAGGTGCTGACATCTTCAAGCTCT 194
QY	988 GTGTAACATCTCCTCTTTGGGCTTGCTGCTGATTCAGTGTGTCATGCTCTCAGCTGAC 1047
Db	195 GACTTCGTGTCTCTTTTGGATTGGGCGCTGTTCCGCGCGGTGATGTCTCTCGCGGAC 254
QY	1048 TCGTCATCTCTGTCGGAGAGTTCATGTTTGCTCGGAATATCTACACAGCTTTCCTTACA 1107
Db	255 TCCTCGGTGCTCTCCGCGCTCCATGTTCGCTCGGAACGTGTACAAATATGATTTTCGT 314
QY	1108 CAAATATGTCAT 1117
Db	315 CAGAAAGCGT 324
RESULT 4	
US-09-270-767-17002	
Sequence 17002, Application US/09270767	
Patent No. 6703491	
GENERAL INFORMATION:	
APPLICANT: Homuser et al.	
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster	
FILE REFERENCE: File Reference: 7346-094	
CURRENT APPLICATION NUMBER: US/09/270,767	
CURRENT FILING DATE: 1999-03-17	
NUMBER OF SEQ ID NOS: 62517	
SOFTWARE: PatentIn Ver. 2.0	
SEQ ID NO 17002	
LENGTH: 324	
TYPE: DNA	
ORGANISM: Drosophila melanogaster	
US-09-270-767-17002	
Query Match	8.0% ; Score 138.8 ; DB 3 ; Length 324 ;
Best Local Similarity	65.5% ; Pred. No. 5.1e-35 ;
Matches 203 ; Conservative 0 ; Mismatches 107 ; Indels 0 ; Gaps 0 ;	
QY	808 GCTCAGGACTGTCCTCTCTCGGACAGCTTTGGGTCCTGTGATGAGCTCTACCCGCCATA 867
Db	15 GCCCAGCTTCGTCTCTATGTTGACCGCGGAGATGATTTTGAATGGCCATTTCCCGGTTG 74

QY	868	TGCATTAGAGCTATTGGAGCTTTCACACAGCTGGAAACGAGCTGCGCTACGGGATTCAGAT	927
Db	75	CTCATCGGAGCGATTGGCCAGGCTTACACCTTGGAAACGAGACAGATTTCAGAGGACCCTAT	134
QY	928	CCCAAGACTAAGAGAGGAAGCAGACATGATTTCTCCGATCGTTTTCGAGTACCTTCGCCCT	987
Db	135	CCCCGTGACCGGAGCGAGACAGACAGATGATTTCTGCCCATGGTGCCTGACATACCTCACGCCCT	194
QY	988	GTGTACATCTCTCTCTTGTGGGCTTGAGTCTGTTTCAGCTGCTGTCATGTCCTCCAGCTAC	1047
Db	195	GACTTCGTGTCTCTTCTTTGGATTGGGCGCTGTTTCCGCCCGCGATGTCTCTCCGCCGAC	254
QY	1048	TCTGTCCATCTCTGTGCGCGAGTTCTATATGTTTGCTCGGAATATCTACAGACTTTCCTTACA	1107
Db	255	TCTCTGGGCTCTCTCCGCCGCTCCATGTTCCGCTCGGAACGTGTACAAATGATTTTCCGT	314
QY	1108	CAAAATGCAT	1117
Db	315	CAGAAAGCGT	324

```

RESULT 5
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIFFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14
Query Match 2.6%, Score 45.8, DB 2, Length 7218
Best Local Similarity 6.8%; Pred. No. 0.0028;

```

	Matches	26;	Conservative	195;	Mismatches	162;	Indels	0;	Gaps	0;
Qy	956	TTTCCCGAGTCGTTCTGACGTAACCTCGCCCTGTGATCATCTCTCTTTGGGCTGGTG	1015							
Db	1106	TT	1165							
Qy	1016	CTGTTTAGCTGCTGTCATGTCCTCACTGACCTGTCGTCATCTGTCGGCGATTTCTATGT	1075							
Db	1166	TT	1225							
Qy	1076	TTTGCTCGAATATCTACAGCTTTCCTTCAGACAAAATGCATCAGACAGAAGAAATTGTCT	1135							
Db	1226	TT	1285							
Qy	1136	GGGTCATGAGGATCAGTCGTCTGTGTGGAGATCTGCACAGCATGGCTTTGCTGA	1195							
Db	1286	TT	1345							
Qy	1196	CGAAGACTGTGATGAGGCTCTGATCCTGACCTGATCTGATCATCATCATCTTCC	1255							
Db	1346	TT	1405							
Qy	1256	CACAGCTGCTCTGTGATCTTTCATCAAAGAACCAACTTATGGGCAGTGCTGTT	1315							
Db	1406	TT	1465							
Qy	1316	ATATTTTGGACATATTCCTGAGA	1338							
Db	1466	GCATGATAGGTAATTACAGTGA	1488							

```

RESULT 6
US-09-543-681A-822
; Sequence 822, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 822
; LENGTH: 1647
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-822

Query Match          2.3%; Score 40; DB 3; Length 1647;
Best Local Similarity 57.0%; Pred. No. 0.079;
Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY      466 ATCATGATGTGGAGTGTGAACATATGGGTCATGTCTCTGCACCTCATTTGCATTTCTTAT 525
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      484 ATCTTGGTGGCCAGTATGATGCAATTATTTGGTGGCACTATTTGGGTTATCTAC 543
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      526 ACCCTAGTGGGGGCTCTACTCTGTGGCATATGATGTTGTGCCAGTATTTCTGCATT 585
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      544 TCGCTTAGCGGACCTTCACCGGTTGCATGACGAGATTTGTACAGTTTCTTCTCG 603
         ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      586 TTATATAGG 593
         |||||
DB      604 ATCTCTGG 611
         |||||

RESULT 7
US-09-662-254B-26
; Sequence 26, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:

```



```

? APPLICANT: Moyer, Richard W.
? APPLICANT: Li, Yi
? APPLICANT: Bawden, Alison Louise
? TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
? TITLE OF INVENTION: Vertebrate Cells
? TITLE OF INVENTION: Vertebrate Cells
? FILE REFERENCE: UF-221C1XCI
? CURRENT APPLICATION NUMBER: US/09/662,254B
? CURRENT FILING DATE: 2000-09-14
? PRIOR APPLICATION NUMBER: 09/086,651
? PRIOR FILING DATE: 1998-05-29
? PRIOR APPLICATION NUMBER: 60/224,479
? PRIOR FILING DATE: 2000-08-10
? NUMBER OF SEQ ID NOS: 80
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 26
? LENGTH: 50000
? TYPE: DNA
? ORGANISM: Amsacta moorei entomopoxvirus
? US-09-662-254B-26

```

Query Match	2.2%;	Score 38.8;	DB 3;	Length 50000;
Best Local Similarity	50.5%;	Pred. No. 2.4;		
Matches 94;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;

QY	1451	TGGTACCTCATCTTTCACCAACATTGGTCTTTCTTA	CTAGGCAAGATCATCTTTGAAA	1510
Db	46550	TTGATACATATATTCCTAATATAAATATGATTTATTA	TTTGGAAAATTCGATATTTTAAAG	46605
QY	1511	GTCGAACTTCCTCCAAAATTAAGATGATTTTGA	TGCTGTGTGCGCAAGGACAGTGAA	1570
Db	46610	GTCATTCATTTTAAAAATAAAGTAAATTCGATATGA	AAATTTGAATTTGATGGTGAAA	46665
QY	1571	AGACATGCGCAAGACCATTCCTGTGTCGAAATG	AAAAATATCAATTTAATGAACCTTGAC	1630
Db	46670	TTATTTTAAATAATGCCAATACATAAAAAATAT	TTATGATCGAGATAAATAAAGTTGGAC	46725
QY	1631	CTGTGCA	1636	
Db	46730	CAGTTAA	46735	

```

RESULT 8
US-09-328-352-2245
; Sequence 2245, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2245
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2245

```

Query Match	2.2%;	Score 38.2;	DB 3;	Length 1506;
Best Local Similarity	50.8%;	Pred. No. 0.29;		

1009 CTTGGTGGTTACGCTGTCATGTCTCAGCTGACTGTCATCCGTGCGGCACT 1068

1066 CTAGCTGATTTTACCTGGGGTTATGAGTACATTAAAGCTGCAGCTTTTGGTATGTTCA 1065

1066 AGTGCACCTAACTGAAGATTGTGACAAAGGCTTCATTGCTAATAATGCATCTCTCAAAAAGAG 1125

1129 ATGTGTGGGTCATGAGGATCTGCTGTGTTGGAGCATCTGGCAAGCATGGC 1187

Db 1126 CTTGTATGGGTTGGCGTATCATGGTGCTTGCAATTGCCGTTCTAGCAATTGTGCTTGC 1184

```

RESULT 9
US-09-640-198D-3
: Sequence 3: Application US/09640198D
: Patent No. 658611
:
: GENERAL INFORMATION:
: APPLICANT: Russell, Stephen
: APPLICANT: Kay Whye, Peng
: TITLE OF INVENTION: System for Monitoring the Location of
: TITLE OF INVENTION: Transgenes
: FILE REFERENCE: 07039-295001
: CURRENT APPLICATION NUMBER: US/09/640,198D
: CURRENT FILING DATE: 2000-08-16
: PRIORITY FILING DATE: 1999-08-17
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 1857
: TYPE: DNA
: ORGANISM: Rattus sp.
: US-09-640-198D-3

```

Query Match	2.2%	Score 37.8;	DB 3;	Length 1857;
Best Local Similarity	49.3%;	Pred. No. 0.46;		
Matches 99; Conservative	0;	Mismatches 102;	Indels 0;	Gaps 0;

Qy	40	TTATGGGAAGATGTTCTCGGGCTGGAGCAATTTTCTGCAATTGAGGGCCACATCAGC	463
Db	413	TGATGGCCACATGCTGTATACAGGATCGATCTACGCCCGCCCTATCCTGAAC	472
Qy	464	TGATCATTTGATGGATGGAACATATGGCATTTGTCTGCACTATTCACATCTTT	523
Db	473	AACTACCGGGTGGACATCTGGGATGCTCTGTCCAGACAGATATATGCACTTGT	532
Qy	524	ATACCTTAATGGGTGGGCTTACTCTGTGGCATATACTGATGTTGTCCAGTATTTGCA	583
Db	533	ACACTACCGTGGTGTGATGAAGCGCGTGTCTGGACAGATGTTCTCCAGGTGTGTAA	592

[illegible]

RESULT 10  
US-08-595-553A-1  
Sequence 1, Application US/08595553A  
Patent No. 6391579  
GENERAL INFORMATION:  
APPLICANT: NANCY CARRASCO, ET AL.  
TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND  
TITLE OF INVENTION: NUCLEIC ACID ENCODING SAME  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN  
STREET: 90 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE  
MEDIUM TYPE: DISKETTE  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/595.553A  
FILING DATE: FEBRUARY 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG J. ARNOLD

REGISTRATION NUMBER: 34,287  
REFERENCE/DOCKET NUMBER: 96700/393  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2839  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: OLIGONUCLEOTIDE  
DESCRIPTION: NO  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: RAT  
INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER  
US-08-595-553A-1

Query Match 2.2%; Score 37.8; DB 3; Length 2839;  
Best Local Similarity 49.3%; Pred. No. 0.63;  
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 404 TGATGGAGAAGATGTTCTGGGGCTGCAGCAATTTCTCTGCATTTAGGGGCCACCATCAGCG 463  
DB 522 TGGTGGCCACAGTGTGTATACAGGCATCGATCTACGCCCTGCCTCATCTGAACC 581  
QY 464 TGATCATGTGATGTGATGGAACATATGGTCATGTCTCTGCACATCGCATTTCTTT 523  
DB 582 AAGTACCGGGTGGACATCTGGGATGCTCTCTGCACAGAAATCATCTGCACCTTGT 641  
QY 524 ATACCTAGTGGGCTCTACTCTGTGCAATATAGATGTGTGTCAGCTATTCTGCA 583  
DB 642 ACACACCGTGGTGTATGAAGCGCGTGTCTGCACAGATGTGTTCCAGGTTGTGTTA 701  
QY 584 TTTTATAGACTGTGATCA 604  
DB 702 TGCTGTGGCTTCTGGGTGA 722

## RESULT 11

US-09-995-007-1  
Sequence 1, Application US/09995007  
Patent No. 6803199  
GENERAL INFORMATION:  
APPLICANT: NANCY CARRASCO, ET AL.  
TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN  
STREET: 90 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE  
DISKETTE  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/995,007  
FILING DATE: 26-NO. 6803199-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/595,553A  
FILING DATE: FEBRUARY 1, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG J. ARNOLD  
REGISTRATION NUMBER: 34,287

REFERENCE/DOCKET NUMBER: 96700/393  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2839  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: <Unknown>  
DESCRIPTION: OLIGONUCLEOTIDE  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: RAT  
INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-995-007-1

Query Match 2.2%; Score 37.8; DB 3; Length 2839;  
Best Local Similarity 49.3%; Pred. No. 0.63;  
Matches 99; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 404 TGATGGAGAAGATGTTCTGGGGCTGCAGCAATTTCTCTGCATTTAGGGGCCACCATCAGCG 463  
DB 522 TGGTGGCCACAGTGTGTATACAGGCATCGATCTACGCCCTGCCTCATCTGAACC 581  
QY 464 TGATCATGTGATGTGATGGAACATATGGTCATGTCTCTGCACATCGCATTTCTTT 523  
DB 582 AAGTACCGGGTGGACATCTGGGATGCTCTCTGCACAGAAATCATCTGCACCTTGT 641  
QY 524 ATACCTAGTGGGCTCTACTCTGTGCAATATAGATGTGTGTCAGCTATTCTGCA 583  
DB 642 ACACACCGTGGTGTATGAAGCGCGTGTCTGCACAGATGTGTTCCAGGTTGTGTTA 701  
QY 584 TTTTATAGACTGTGATCA 604  
DB 702 TGCTGTGGCTTCTGGGTGA 722

## RESULT 12

US-09-949-016-15637/C  
Sequence 15637, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15637  
LENGTH: 48794  
TYPE: DNA  
ORGANISM: Human  
FEATURES:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(48794)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15637

Query Match 2.2%; Score 37.8; DB 3; Length 48794;  
Best Local Similarity 55.8%; Pred. No. 5.1;

```
Matches 72; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 1414 ATATAATCAGAGGTTCCCATTTAAACCTCTCCATGGTATCCATCTTTTACCAC 1473
    |||||
Db 22938 ATATACATCTCTCTCTAGTTAAACCTATCTTTATCTACATCTTTAGACC 22879
QY 1474 ATTTGTTCTTATCTAGCCAGATCTATTTGAAAGTGAACCTTGCCCAATTA 1533
    |||||
Db 22878 CTAATTCACAGATCTTTGCCATTTATGCTCTTAAACCTCGCCCTTACCTACATTA 22819
QY 1534 GATGATTT 1542
    |||||
Db 22818 TTGGTATCT 22810

RESULT 13
US-09-513-999C-2240/c
; Sequence 2240, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclercq, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2240
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 29..478
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 141
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 38
; OTHER INFORMATION: Xaa=Leu or Pro
; US-09-513-999C-2240

Query Match 2.1%; Score 37; DB 3; Length 478;
Best Local Similarity 64.7%; Pred. No. 0.31; 30; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 30;

QY 1177 ACAGCCATGCTTGTGTAAGAGCTGTATGGCTGTGACCTGACCTT 1236
    |||||
Db 91 AGAGCCTTGCTTCAGTCGAGGTTATTTAGGCGCTCTTCCCTGAGCTC 32
QY 1237 GCTTACATATCATCTTCCACAGC 1261
    |||||
Db 31 CATTTCCGATCAGATTCTCCAGC 7

RESULT 14
US-09-539-333D-1/c
; Sequence 1, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marita
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essieux, Laurent

; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET 047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..1107
; OTHER INFORMATION: 5'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1108..1289
; OTHER INFORMATION: exon A g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14877..14920
; OTHER INFORMATION: exon B g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18778..18862
; OTHER INFORMATION: exon Bbis g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25593..25740
; OTHER INFORMATION: exon C g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29388..29502
; OTHER INFORMATION: exon D g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 29967..30282
; OTHER INFORMATION: exon E g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 64666..64812
; OTHER INFORMATION: exon F g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 65505..65853
; OTHER INFORMATION: exon G g35018 gene
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 65854..67854
; OTHER INFORMATION: 3'regulatory region g35018 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 94124..94964
; OTHER INFORMATION: exon g35017
; FEATURE:
```

NAME/KEY: exon  
LOCATION: 201188..201234  
OTHER INFORMATION: exon S g35030 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 214676..214793  
OTHER INFORMATION: exon T g35030 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 215702..215746  
OTHER INFORMATION: exon U g35030 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 216836..216915  
OTHER INFORMATION: exon V g35030 gene  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 213818..215818  
OTHER INFORMATION: 3' regulatory region g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 215819..215941  
OTHER INFORMATION: exon R complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 215819..215975  
OTHER INFORMATION: exon Rbis complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 216661..216952  
OTHER INFORMATION: exon Qbis complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 217027..217061  
OTHER INFORMATION: exon Q complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 229647..229742  
OTHER INFORMATION: exon X complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 230408..230721  
OTHER INFORMATION: exon P complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 231272..231412  
OTHER INFORMATION: exon Obis complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 231787..231880  
OTHER INFORMATION: exon O2 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 231870..231879  
OTHER INFORMATION: exon O1 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 234174..234321  
OTHER INFORMATION: exon O complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 237406..237428  
OTHER INFORMATION: exon Nbis complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 239719..239807  
OTHER INFORMATION: exon N2 complement g34872 gene  
FEATURE:  
NAME/KEY: exon

LOCATION: 239719..239853  
OTHER INFORMATION: exon N complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240569  
OTHER INFORMATION: exon M1117 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240596  
OTHER INFORMATION: exon M1090 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240617  
OTHER INFORMATION: exon M1069 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240644  
OTHER INFORMATION: exon MS2 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240824  
OTHER INFORMATION: exon M862 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..240994  
OTHER INFORMATION: exon M692 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240528..241685  
OTHER INFORMATION: exon M1 complement g34872 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 240800..240993  
OTHER INFORMATION: exon M51 complement g34872 gene  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 241686..243685  
OTHER INFORMATION: 5' regulatory region g34872 gene  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 290652..292652  
OTHER INFORMATION: 3' regulatory region g34665 gene  
FEATURE:  
NAME/KEY: exon  
LOCATION: 292653..292841

Query Match 2.1%; Score 37; DB 3; Length 319608;  
Best Local Similarity 47.3%; Pred. No. 37;  
Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1418 ACAATCAAGAGTTCCCATTTAAACTCTCTCCATGTTACCTCATTTCTTACCAACATTT 1477  
DB 156436 ACATCTGATGTTCTCAGTCAAAACAAATTTCCACAAATTCCTTAACTGCTGAGAAAGACA 156377  
QY 1478 GTGTTCTTATCTACCAAGTATCTATTGAAAGTGAACCTTGCTCCAAATTAAGATG 1537  
DB 156376 GAATTCATCTATTAAACATGATGCTTCACTGACGTCAATGATTTTTCACATTTTG 156317  
QY 1538 TATTTGATGCTGTTGCTCCCAAGGACAGTGAAGAGACATGACACCAATTCATGTC 1597  
DB 156316 TCAATTAATTTCTGATATCATGATTAACAGCAAAAGAAATTTAATACTGCTGGCCACTGGC 156257  
QY 1598 GAATGAATAATATCAATTAATGAACCTTGCACTGTGAAACCTGCGCAGAGCCTTA 1654  
DB 156256 TTAAAAAATTTTGAAGAAAAATGAACATGCAAAATTAATGATGATTTAGAGTCCA 156200

RESULT 15  
US-09-679-409-1/c  
; Sequence 1, Application US/09679409  
; Patent No. 6555316  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENE, PROTEINS AND BIALLELIC MARKERS  
FILE REFERENCE: 53,US15.CIP  
CURRENT APPLICATION NUMBER: US/09/679,409  
CURRENT FILING DATE: 2000-10-03  
PRIORITY APPLICATION NUMBER: 09/539,333  
PRIORITY FILING DATE: 2000-03-03  
PRIORITY APPLICATION NUMBER: 09/416,384  
PRIORITY FILING DATE: 1999-10-12  
PRIORITY APPLICATION NUMBER: 60/168,088  
PRIORITY FILING DATE: 1999-11-30  
NUMBER OF SEQ ID NOS: 134  
SOFTWARE: Patent.pm  
SEQ ID NO 1  
LENGTH: 319608  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 199122..201122  
OTHER INFORMATION: 5'regulatory region  
NAME/KEY: exon  
LOCATION: 201123..201234  
OTHER INFORMATION: exon S  
NAME/KEY: exon  
LOCATION: 201123..201560  
OTHER INFORMATION: exon S2  
NAME/KEY: exon  
LOCATION: 214676..214793  
OTHER INFORMATION: exon T  
NAME/KEY: exon  
LOCATION: 215702..215746  
OTHER INFORMATION: exon U  
NAME/KEY: exon  
LOCATION: 216836..216994  
OTHER INFORMATION: exon V  
NAME/KEY: exon  
LOCATION: 216836..217077  
OTHER INFORMATION: exon V2  
NAME/KEY: exon  
LOCATION: 217671..217764  
OTHER INFORMATION: exon V1  
NAME/KEY: exon  
LOCATION: 227655..227736  
OTHER INFORMATION: exon V4  
NAME/KEY: exon  
LOCATION: 238715..238919  
OTHER INFORMATION: exon V3  
NAME/KEY: exon  
LOCATION: 240440..240673  
OTHER INFORMATION: exon W  
NAME/KEY: exon  
LOCATION: 240440..241153  
OTHER INFORMATION: exon W2  
NAME/KEY: exon  
LOCATION: 241072..241291  
OTHER INFORMATION: exon X  
NAME/KEY: exon  
LOCATION: 244353..244561  
OTHER INFORMATION: exon Y  
NAME/KEY: exon  
LOCATION: 246273..247802  
OTHER INFORMATION: exon Z  
NAME/KEY: misc\_feature  
LOCATION: 247803..249803  
OTHER INFORMATION: 3'regulatory region  
NAME/KEY: allele  
LOCATION: 8316  
OTHER INFORMATION: 99-27943-150 : polymorphic base G or C  
NAME/KEY: allele

LOCATION: 21672  
OTHER INFORMATION: 99-27935-193 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 65485  
OTHER INFORMATION: 8-128-33 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 95396  
OTHER INFORMATION: 99-31960-363 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 107281  
OTHER INFORMATION: 99-24656-260 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 160640  
OTHER INFORMATION: 99-24639-163 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 160876  
OTHER INFORMATION: 99-24634-108 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 168974  
OTHER INFORMATION: 99-7652-162 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 170810  
OTHER INFORMATION: 99-16100-147 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 173358  
OTHER INFORMATION: 99-5862-167 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 189957  
OTHER INFORMATION: 99-5919-215 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 197163  
OTHER INFORMATION: 99-24658-410 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 200778  
OTHER INFORMATION: 8-303-235 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 202651  
OTHER INFORMATION: 8-300-221 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 202679  
OTHER INFORMATION: 8-300-193 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 203378  
OTHER INFORMATION: 8-299-128 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204138  
OTHER INFORMATION: 8-296-213 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 204605  
OTHER INFORMATION: 8-252-190 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 204934  
OTHER INFORMATION: 99-24644-194 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 205206  
OTHER INFORMATION: 8-295-248 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 205329  
OTHER INFORMATION: 8-295-125 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 206064  
OTHER INFORMATION: 8-293-130 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 206545  
OTHER INFORMATION: 8-292-198 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 207313  
OTHER INFORMATION: 8-251-322 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 208285  
OTHER INFORMATION: 8-289-322 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 208960

Job time : 352 secs

```
OTHER INFORMATION: 8-287-249 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 209123
OTHER INFORMATION: 8-287-86 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 209631
OTHER INFORMATION: 8-285-319 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210361
OTHER INFORMATION: 8-283-278 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210463
OTHER INFORMATION: 8-283-176 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 210486
OTHER INFORMATION: 8-283-153 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210583
OTHER INFORMATION: 8-283-56 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 210879
OTHER INFORMATION: 8-282-345 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 210979
OTHER INFORMATION: 8-282-245 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 211050
OTHER INFORMATION: 8-282-174 : variable motif AAAGG or GAAGGAAGGAAGGAAGA
NAME/KEY: allele
LOCATION: 211132
OTHER INFORMATION: 8-282-92 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 211247
OTHER INFORMATION: 8-281-367 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211315
OTHER INFORMATION: 8-281-299 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 211366
OTHER INFORMATION: 8-281-248 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 212520
OTHER INFORMATION: 8-279-197 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 212821
OTHER INFORMATION: 8-278-289 : polymorphic base C or T
NAME/KEY: allele
```

Query Match 2.1%; Score 37; DB 3; Length 319608;

Best Local Similarity 47.3%; Pred.No. 37;

Matches 112; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
OY 1418 ACAATCAGAGTTCCCATTTAAACTCTCCATGTTACTTACCTCATTTCTTACCAACATTT 1477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156436 ACATCTGATGTTCTCACCAAAACAAATTTCCCAATTCCTACTGCTGATGAACGACA 156377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1478 GTGTTTCTTATCTAGCCAAAGTATCTATTGAAAGTGAACCTTGCCCTCCAAATTAGATG 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156376 GAATTCCTACTATTAAACAATGATCACTTTCACCTGACGTAATGATTTTCCATTTTGG 156317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1538 TATTGATGCTGTGTGCGACAGGACAGTGAAGAAGAACATGACAAAGACCATTCATGTC 1597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156316 TCAATTAATTCGTGTATCATGATTAACAGCAAGAAATTTAATTAATGCTGCGCACTGSC 156257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1598 GAAATGAAATATTCAAATTAATGAACCTGACCTGTGAACCTGGCGAGACCTTA 1654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156256 TTAATAAATTTTAAAGAAAAATGAACATGCAATTAATTAATTAATGATGATCCAA 156200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 10, 2006, 16:37:10

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: July 10, 2006, 14:09:32 ; Search time 312 Seconds  
(without alignments)

5217.515 Million cell updates/sec

Title: US-10-724-806-4

Perfect score: 2993

Sequence: 1 MPFHEGVAVILFLYLFL.....EALLDVDSPEGGSTEDNLQ 580

Scoring table: XGAPSUM2

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 segs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.model -DEV=xlh  
-O=/abs/ABSSWEB\_spool/US10724806/runat.10072006.140252.7805/app\_query.fasta.1  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abss07 -USER=US10724806 -CCEN 1.1.204 -runat.10072006.140252.7805 -NCPu=6  
-ICPU=3 -NO\_MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA.\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1-COMB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5-COMB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A-COMB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B-COMB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7-COMB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H-COMB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCIDS-COMB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP-COMB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE-COMB.seq.\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backflist1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2795	93.4	1743	3	US-09-657-252-1
2	1099.5	36.7	1461	5	US-09-974-300-501
3	351	11.7	324	3	US-09-270-767-1720
4	351	11.7	324	3	US-09-270-767-17002
5	318	10.6	2475	3	US-09-799-451-932
6	310	10.4	2028	3	US-10-162-012-28
7	310	10.4	2326	3	US-10-162-012-26
8	307	10.3	2107	3	US-09-799-451-234

9	298	10.0	1566	3	US-09-540-236-273	Sequence 273, App
10	298	10.0	6703	3	US-09-596-002-7	Sequence 7, Appl1
11	286	9.9	2450	3	US-09-949-016-5012	Sequence 5012, Ap
12	286	9.9	3342	3	US-09-252-991A-7814	Sequence 7814, Ap
13	293.5	9.8	1455	3	US-09-252-991A-7678	Sequence 7678, Ap
14	293.5	9.8	1485	3	US-09-252-991A-7528	Sequence 7528, Ap
15	284	9.5	2238	2	US-07-841-651-1	Sequence 1, Appl1
16	276.5	9.2	1647	3	US-09-543-681A-822	Sequence 822, App
17	274.5	9.2	2394	3	US-09-799-451-931	Sequence 931, App
18	271.5	9.1	2157	3	US-09-657-960-2	Sequence 2, Appl1
19	271.5	9.1	10302	3	US-09-657-960-1	Sequence 1, Appl1
20	271.5	9.1	25356	3	US-09-976-594-750	Sequence 750, App
21	265	8.9	1830121	3	US-09-557-884-1	Sequence 1, Appl1
22	265	8.9	1830121	3	US-09-643-990A-1	Sequence 1, Appl1
23	265	8.9	1830121	3	US-10-158-865-1	Sequence 1, Appl1
24	260.5	8.7	1695	3	US-09-602-787A-495	Sequence 495, App
25	259	8.7	1527	3	US-09-489-039A-370	Sequence 370, App
26	258	8.6	3280	3	US-09-799-451-818	Sequence 818, App
27	254	8.5	1506	3	US-09-328-352-2245	Sequence 2245, Ap
28	245.5	8.2	1557	3	US-09-134-001C-1907	Sequence 1907, Ap
29	244.5	8.2	987	5	US-09-974-300-500	Sequence 500, App
30	243	8.1	1776	3	US-09-540-236-570	Sequence 570, App
31	241.5	8.1	1332	3	US-09-252-991A-10993	Sequence 10993, A
32	241.5	8.1	1656	3	US-09-252-991A-11258	Sequence 11258, A
33	241.5	8.1	2217	3	US-09-252-991A-11190	Sequence 11190, A
34	241	8.1	89047	3	US-09-596-002-34	Sequence 34, Appl1
35	240.5	8.0	1491	3	US-09-489-039A-1601	Sequence 1601, Ap
36	236.5	7.9	1665	3	US-09-543-681A-12931	Sequence 12931, A
37	233.5	7.8	1551	3	US-09-252-991A-12931	Sequence 12931, A
38	233.5	7.8	1659	3	US-09-252-991A-13081	Sequence 13081, A
39	233.5	7.8	2073	3	US-09-252-991A-12858	Sequence 12858, A
40	232	7.8	13542	3	US-08-956-171E-154	Sequence 154, App
41	232	7.8	13542	3	US-08-956-171E-154	Sequence 154, App
42	230.5	7.7	1521	3	US-09-543-681A-4074	Sequence 4074, Ap
43	230	7.7	1698	3	US-09-489-039A-1243	Sequence 1243, Ap
44	226.5	7.6	1461	3	US-09-543-681A-2066	Sequence 2066, Ap
45	224.5	7.5	1830121	3	US-09-557-884-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-657-252-1  
Sequence 1, Application US/09657252  
Patent No. 6500643  
GENERAL INFORMATION:  
APPLICANT: Wu, Dong-Hai  
APPLICANT: Gu, Yunrong  
APPLICANT: Millard, William  
APPLICANT: He, Yun-Je  
TITLE OF INVENTION: Human High Affinity Choline Transporter cDNA  
FILE REFERENCE: M8B00-639  
CURRENT APPLICATION NUMBER: US/09/657,252  
CURRENT FILING DATE: 2000-09-07  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 1743  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1743)  
US-09-657-252-1

Alignment Scores:  
Pred. No.: 3,42e-314  
Score: 2795.00  
Percent Similarity: 96.6%  
Best Local Similarity: 92.6%  
Query Match: 93.4%  
Length: 1743  
Matches: 537  
Conservative: 23  
Mismatch: 20  
Indels: 0  
Gaps: 0

US-10-724-806-4 (1-580) x US-09-657-252-1 (1-1743)

QY 1 MetProPheHisValGluGlyLeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeu 20  
Db 1 ATGGCTTCCATGCTGGAAGGAGCTGATGAGTATCATCGTGTCTTACCTTATTTTGGCG 60

QY 21 ValGlyIleTPRAlaAlaTPrLyThrLyAsnSerGlyAsnProGluIAsnSerGlu 40  
Db 61 GTTGAATAATGGCGCTGCGTGGAGAACCAAAAACAGTGGCAGCGAGAAAGCGCGCA 120

QY 41 AlaIleIleValGlyGlyYArgAspIleGlyLeuLeuValGlyGlyPheThrMetThrAla 60  
Db 121 GCCATCATAGTGGTGGCGCGAGATATGTTATTTGGTGGATTTACCATGACAGCT 180

QY 61 ThrTPRValGlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCys 80  
Db 181 ACCTGGGTCGAGAGGAGGATATCATGACAGCTGAGAGAGAGTTATGATACAGGTTAT 240

QY 81 GlyLeuAlaTPRAlaHisAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeu 100  
Db 241 GGCCTAGCTGGGCTCAGGACCAATGGATATCTCTTAGCTGATATTTAGTGGCGCTG 300

QY 101 PhePheAlaLysProMetArgSerLysGlyTyrValIleThrMetLysAspProPheLysGln 120  
Db 301 TTCTTTGCAAAACCTATGCGTTCAAAAGGGGATATGACCATTTAGACCCGTTTCAAGCA 360

QY 121 IleTyrGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLysMetPhe 140  
Db 361 ATCTATGAGAAACGATGGGCGGAGCTCCGTTTATCTCGACAGCTGATGGAGAAATGTTTC 420

QY 141 TPRAAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160  
Db 421 TGGGCTGAGCAATTTTCTGCTTTGGAGCCACCATCAGCGTATCATATGATGGAT 480

QY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyrThrLeuValGlyGly 180  
Db 481 ATGCACATTTCTGTATCATCTCTGACATCATTTGCCACTCTGTACACTGGTGGAGGG 540

QY 181 LeuTyrSerValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTyr 200  
Db 541 CTCATATCTGTGGCTGACAGCTGATGCGTTCAGCTCTTTGCAATTTTGTAGGGCTGTGG 600

QY 201 IleSerValProPheAlaLeuSerHisProAlaValIleAspIleGlyLeuThrAlaVal 220  
Db 601 ATGACCGCTCCCTTTCATGTCATCTGACATCTGACATCGGAGATCGGCTTCACTCGTGG 660

QY 221 HisAlaLysTyrGlnSerProTyrLeuGlyThrIleGlnSerValGlnValTyrThrTyr 240  
Db 661 CATGCCAAATACCAAAAGCGCTGGCTGGGAACCTGTGATCTCATGTAGTCTTCTTGG 720

QY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyrGlnAlaTyrPheGlnArg 260  
Db 721 CTGTAGATGTTTCTGTGTTGATGCTGGGGTGAATCCATGCAAGCATATTTCAAGAGG 780

QY 261 ValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIlePheGly 280  
Db 781 GTTCTCTCTCTCTCTCAGCCACCTATGCTCAAGTGTGCTCTCTCTGACACTTTCGGG 840

QY 281 CysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTyr 300  
Db 841 TGCCGTGGTGGATGGCCATCCCAACCATCTATTTGGGCCATTTGGACCTCCACAGACTGG 900

QY 301 AsnGlnThrAlaTyrGlyTyrProAspProLysThrLysGlnGluAlaAspMetIleLeu 320  
Db 901 AACGAGACTGATATGGGCTTCCAGATCCCAAGACTACAGAAAGAGCAGACATGATTTTA 960

QY 321 ProIleValLeuGlnTyrLeuCysProValTyrIleSerPhePheGlyLeuGlyAlaVal 340  
Db 961 CCAATGTTCTGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 341 SerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAla 360  
Db 1021 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

QY 361 ArgAsnIleTyrGlnLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTyrVal 380  
Db 1081 CGGAACATCTACAGCTTTCCTTCAGACAAATGCTTGGACAAAGAAATGCTTGGGTT 1140

QY 381 MetArgIleThrAlaLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLys 400  
Db 1141 ATGCAATCACAGTGTGTTGTGGACATCTGCAACAGCCATGCTGCTGTCAGAAA 1200

QY 401 ThrValTyrGlyLeuTyrTyrLeuSerSerAspLeuValTyrIleIleIlePheProGln 420  
Db 1201 ACTGTATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

QY 421 LeuLeuCysValLeuPheIleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIle 440  
Db 1261 CTGCTTGTGTACTCTTGTGTAGGAAACCAACCTATGGGCGCTGCGAGGTTATGTT 1320

QY 441 PheGlyLeuPheLeuArgIleThrGlyGlyLysProTyrLeuTyrLeuGlnProLeuIle 460  
Db 1321 TCTGGCTCTTCTGAGATACTGAGAGGAGCCATATGTTATCTTTCAGCCCTTGATTC 1380

QY 461 PheTyrProGlyTyrTyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLys 480  
Db 1381 TTCTACCTGCTGCTATTTACCTGATGATATGATATATATATATCAAAATTTCCATTTTAA 1440

QY 481 ThrLeuSerMetValIleThrSerPheThrAsnIleCysValSerTyrLeuAlaLysTyr 500  
Db 1441 ACACCTGCAAGTGTATCATCTTTTAAACCAACATTTGCACTCTCTATCTGCAAGTAT 1500

QY 501 LeuPheGlnSerGlyTyrLeuProProLysLeuAspValPheAspAlaValAlaArg 520  
Db 1501 CTATTTGAAAGGAACTTTCGACCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560

QY 521 HisSerGlnGluAsnMetAspLysThrIleLeuValArgAsnGluAsnIleLysLeuAsn 540  
Db 1561 CACAGTGAAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620

QY 541 GluLeuAlaProValLysProArgGlnSerLeuThrLeuSerSerThrPheThrAsnLys 560  
Db 1621 GAACCTTGACCTGTGTAACCAACGACAGCATGACCTCAGCTCAACTTTCACCAATTA 1680

QY 561 GluAlaLeuLeuAspValAspSerSerProGlnGlySerGlyTyrGluAspAsnLeuGln 580  
Db 1681 GAGGCTTCTGATGTTGATTCAGTCCAGAGGCTCTGGAGCTGAAGATTAATTTACA 1740

RESULT 2  
US-09-974-300-501  
Sequence 501, Application US/09974300  
Patent No. 7018784  
GENERAL INFORMATION:  
APPLICANT: Berka, Randy M.  
APPLICANT: Clausen, Ib Groth  
TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
FILE REFERENCE: 10085-500-US  
CURRENT APPLICATION NUMBER: US/09/974,300  
PRIOR FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 09/680,598  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/279,526  
NUMBER OF SEQ ID NOS: 8481  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 501  
LENGTH: 1461  
TYPE: DNA  
ORGANISM: Bacillus licheniformis  
US-09-974-300-501

Alignment Scores:  
Pred. No.: 2,63e-117 Length: 1461  
Score: 1099.50 Matches: 232  
Percent Similarity: 62.84 Conservative: 89



Best Local Similarity: 45.4% Mismatches: 147  
 Query Match: 36.7% Indels: 43  
 DB: 5 Gaps: 8

US-10-724-806-4 (1-580) x US-09-974-300-501 (1-1461)

QY 4 HisValGlyLeuValAlaIleLeuPheTyrLeuLeuIlePheLeuValGlyIle 23  
 DB 2 CATGGGGGGGGCTTATCTCCATGATGATATCTTGGCTTACCCATATTAAGGCGCT 61  
 QY 24 TTPAlaAlaTPrIyAsnSerGlyAsnProGluIuArgSerGluAlaIleIle 43  
 DB 62 TTTTACGCGCAAGAAATCAATCTTTTCCGAT-----ATGATC 103  
 QY 44 ValGlyGlyArgAspIleLeuLeuValGlyGlyPhePheMetThrAlaThrPVal 63  
 DB 104 GTGGCAAAAGAGCATGCCGTTTTCGTGGAGATGCTGACATGGCCGACATGGGCT 163  
 QY 64 GlyGlyGlyTyrIleAsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAla 83  
 DB 164 GCGGGGGGATATTATTAACGGAAGCCGATCGACTTACAGCGAC-----GGCTCATC 217  
 QY 84 TTPAlaAlaAlaProIleGlyTyrSerLeuSerLeuIleLeuGlyGlyLeuPheAla 103  
 DB 218 TGGGCCCAAGCGCTTGGGGCTACGATGAGCTGATTAACGCGGATATTTCTTCCGC 277  
 QY 104 LysProMetArgSerIyGlyTyrValThrMetLeuAspProPheGlyGlyIleTyrGly 123  
 DB 278 AGAAAAATCGCGCATCATTAATTAAGCATTAATGATCTCTCGAACAAGCGCTTCCG 337  
 QY 124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlyLeuPheTPrAlaAla 143  
 DB 338 AGCGGATGGCGGCTTTATATATACGGCGCTTTAGAGAAATGTTTGGACCGCC 397  
 QY 144 AlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAspValAsnIle 163  
 DB 398 GCGATCTTAACGCGCATGGGACAGCACTTTCGAAATGATCTGAAATGATTCGCAACG 457  
 QY 164 SerValIleValSerIleValIleAlaIleLeuTyrThrLeuValGlyGlyLeuTyrSer 183  
 DB 458 TCGATTATTTCTTGGCGAGATGCGCATGCGCTTAACGCTGGCCGAGCGCATGTGGCC 517  
 QY 184 ValAlaTyrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTPrIleSerVal 203  
 DB 518 GTTGCTTACAGATCTCTTCAATGATGATTTGGTGGCGCTTTTATAGCGTC 577  
 QY 204 ProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaValHisAlaLys 223  
 DB 578 CCATTGTATTGTGCAATGTCGCTGCTCAT-----ACGATATGGCAAAAT 625  
 QY 224 Tyr-----GlnSer 226  
 DB 626 TACAGGACGAGATTTCGCGACAGCGCCAACTGCTTCCGCGTTGACGCGGTGAAAC 685  
 QY 227 ProTPrLeuGlyThrIleGlySerValGluValTyrThrTPrLeuAspAsnPheLeu 246  
 DB 686 CCGGATTGGGGAACCTG-----TTTGGAACTGGTGGGATATATGGCTCTC 733  
 QY 247 LeuMetLeuGlyGlyIleProTPrGlnAlaTyrPheGlnArgValLeuSerSerSer 266  
 DB 734 TTAATTTTCGAGAGTTCATGCGAGGTGACTTTCAGGCGCTTCTTCGCGAAATCG 793  
 QY 267 AlaThrTyrAlaGlnAlaLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeu 286  
 DB 794 GAAAGCGCGGCACTTGGCGAGTCGATTAATGCGGAGTGTCTTCATTCGCCGCAT 853  
 QY 287 ProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTPrAsnGlnThrAlaTyrGly 306  
 DB 854 CCGTGGTATATCATCGAGCTGCGGAAACAGTACGATTCGAGCTGTTCGAGGAGC 913  
 QY 307 TyrPTrAspProLysThrIySgIuGluAlaAspMetIleLeuProIleValLeuGlnTyr 326  
 DB 914 GCTCCGAGT-----AACCGGCGATATTTTCCGCAAAAGCTTGGCTAT 958

QY 327 LeuCysProValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaMetSer 346  
 DB 959 TTGAGCCCAAGATCATTCAGAGCCCTCGGCTTGGTGCAATCGCAGACCGCATCTCA 1018  
 QY 347 SerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeu 366  
 DB 1019 AGCATGACTCATCATCTTATCGGCATCATGCGCGCATGGAATTTTACCGTCCG 1078  
 QY 367 SerPheArgGlnAsnAlaSerAspLysGluIleValTPrValMetArgIleThrValLeu 386  
 DB 1079 CTCATCAACCCGAGGCGCAACCAAAACAGCTGCAAAAAGTGTCAAGCTCAATCAT 1138  
 QY 387 ValPheGlyAlaSerAlaThrAlaMetAlaLeuThrLysThrValTyrGlyLeuTPr 406  
 DB 1139 TTGTTGGCGCGGAGACCGGTCATCGGCTGTAATGCAAAAGCTTTATCTTATG 1198  
 QY 407 TyrLeuSerSerAspLeuValTyrIleIleIlePheProGlnLeuLeuCysValLeuPhe 426  
 DB 1199 TATTTGGCTTCGATTTATTTATTCATTTCTTTCCCAAGTTAACATGGCCCTCTT 1258  
 QY 427 IleLysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArg 446  
 DB 1259 TATAAAGAGCAAACTTTACGGGTGATGCTGATTTGCAAGTTGCAGTTCATTCGAG 1318  
 QY 447 IleThrGlyGlyLeuProTPrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyr 466  
 DB 1319 CTCGGCGGTGGTGAACCCGATTCGCGCTTCCGCTTCTGCGCATTCG--ATGATT 1375  
 QY 467 SerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThr 486  
 DB 1376 GAAGACGAGCAGCTCTC-----TTTCTTCCGACGCTGGTGACATGACCA 1423  
 QY 487 SerPhePheThrAsnIleCysValSerTyrLeu 497  
 DB 1424 GCATTTATGACGATCTTCGCTGTGCGAATTG 1456

# RESULT 3

US-09-270-767-1720  
 ; Sequence 1720. Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; NUMBER OF SEQ ID NOS: 62517  
 ; SOFTWARE: Patentin Ver. 2.0  
 ; SEQ ID NO 1720  
 ; LENGTH: 324  
 ; TYPE: DNA  
 ; ORGANISM: Drosophila melanogaster

US-09-270-767-1720

Alignment Scores:  
 Pred. No.: 2 05e-31 Length: 324  
 Score: 351.00 Matches: 69  
 Percent Similarity: 81.3% Conservative: 18  
 Best Local Similarity: 64.5% Mismatches: 20  
 Query Match: 11.7% Indels: 0  
 DB: 3 Gaps: 0

US-10-724-806-4 (1-580) x US-09-270-767-1720 (1-324)

QY 266 SerAlaThrTyrAlaGlnAlaLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAla 285  
 DB 3 ACGGACGAGAAAGGCCAGCTTGTCTATGTGGACGCGCGGAGCATTTTGAAGGCC 62  
 QY 286 LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTPrAsnGlnThrAlaTyr 305  
 DB 63 ATTCCCGGTGCTCATCGAGGAGTTGCCAAGGCTACCTTGGAACAGACAGATTCAC 122

Oy	306	GlyTyrProAspProLysThrIlysgIuAlaAspMetIleuPheProIleValLeuGln	325
Db	123	AAGGACCCCTATCCCTGTGACCGGTGACGACGACGACGATATTCTGCCCATGGTGTGCGAG	182
Oy	326	TyrLeuCySProValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMet	345
Db	183	TACCCACAGCCCTGACTCGTGTCTCTTGTGGATTGGGGCGGTTCGCCGCCGGATG	242
Oy	346	SerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaIlyrGln	365
Db	243	TCCTCCGCCGACTCTCCGGTGTCTTCGCCGCTCATGTTCCTCGGAACGTGTACAA	302
Oy	366	LeuSerPheArxGlnAsnAla	372
Db	303	TTGATTTTCCGTACAGAGCGC	323
RESULT 4			
US-09-270-767-17002			
; Sequence 17002, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 17002			
; LENGTH: 324			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-17002			
Alignment Scores:			
Pred. No.:	2.05e-31	Length:	324
Score:	351.00	Matches:	69
Percent Similarity:	81.3%	Conservative:	18
Best Local Similarity:	64.5%	Mismatches:	20
Query Match:	11.7%	Indels:	0
DB:	3	Gaps:	0
US-10-724-806-4 (1-580) x US-09-270-767-17002 (1-324)			
Oy	266	SerAlaThrTyrAlaGlnValLeuSerPheLeuAlaAlaPheGlyCysIleuAlMetAla	285
Db	3	ACGGCAGGAAAGGCGCCAGCTTCTGTCTATGTTGACGCCCGGAGGCAATTTGATGGCC	62
Oy	286	LeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspIlyrPheAsnIlyrAlaTyr	305
Db	63	ATTCCTCCCGTGTCTATCGGAGGATTGCCAAGGCTACACCTTGGAGACGACGACGATTAC	122
Oy	306	GlyTyrProAspProLysThrIlysgIuAlaAspMetIleuPheProIleValLeuGln	325
Db	123	AAGGACCCCTATCCCTGTGACCGGTGACGACGACGACGATATTCTGCCCATGGTGTGCGAG	182
Oy	326	TyrLeuCySProValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMet	345
Db	183	TACCCACAGCCCTGACTCGTGTCTCTTGTGGATTGGGGCGGTTCGCCGCCGGATG	242
Oy	346	SerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaIlyrGln	365
Db	243	TCCTCCGCCGACTCTCCGGTGTCTTCGCCGCTCATGTTCCTCGGAACGTGTACAA	302
Oy	366	LeuSerPheArxGlnAsnAla	372
Db	303	TTGATTTTCCGTACAGAGCGC	323
RESULT 5			
US-09-799-451-932			
; Sequence 932, Application US/09799451			
; Patent No. 6783969			
; GENERAL INFORMATION:			

[illegible]

Db 800 ATCTACATCTTCACCAAGATCTCGTAGACATGTATGACAGTCCATCTTC 850  
Qy 150 GYAlaThrIleSerValIleIleAspValAspValAsnIleSerValIleValSerAla 169  
Db 851 -----ATCCAGAGCTCTTTCACCTGATCTGTACCTGGCCATAGTGGGCTACTG 901  
Qy 170 LeuIleAlaIleLeuThrIleLeuValGlyGlyLeuTyrSerValAlaTyrThrAspVal 189  
Db 902 GCCATACGCTGTATACACGCTGCTGCTGGCTGGCTGGCTGTGATCTACACGATGCC 961  
Qy 190 ValGlnLeuPheCysIlePheIleGlyLeuTyrIleSerValProPheAlaLeuSerHis 209  
Db 962 CTGCACAGCCTGATCATGCTTATAGAGCGCTCACCTTGATGGGCTAAC-----AGTTTC 1015  
Qy 210 ProAlaValThrAspIleGlyPheThrAlaValAlaHisAlaTyr 224  
Db 1016 GCCGCGGTGGT-----GGGATGGAAGAGCTAAGAGAGATCTTGGCCCTGGCT 1069  
Qy 224 ----- 224  
Db 1070 AGCAACCGAGTAGAAGACAGCAGCTCGGGCTGCCCGGGAAGATGCCCTTCCATTTTC 1129  
Qy 225 -----GlnSerProTyrPheGlyThrIleGlnSerValGluVal 237  
Db 1130 CGAGATCCGCTGACATCTGATCTCCGCTGGCGGGGCTCTATTGGAGATGCCATCCCA 1189  
Qy 238 -----TyrThrTyrPheAspPheLeuLeuLeuMetLeuGlyIleProTyrGln 255  
Db 1190 TCCCTCTGTACTGTGACGAGAT-----CAG 1216  
Qy 256 AlaTyrPheGlnArgValLeuSerSerSerAlaThrTyrAlaGln 271  
Db 1217 GTGATGTCCAGGAGCTGCTGCTGCCAAGACCTGCTCCATCCCAAGAGGCTGCTG 1276  
Qy 272 ValLeuSerPheLeuAlaAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIle 291  
Db 1277 ATGGTGCACTACCTGAAGGCTGCCCTCTTCATATGATGTTCTCT 1324  
Qy 292 GlyAlaIleGlyAlaSerThrAspThrAspGlnThrAlaTyrIleTyrProAspProlys 311  
Db 1325 GGGATGTGACGCGCATCTCTTCCAGATCAGTGGCGCTGCA-----GATCCAGAG 1378  
Qy 312 ThrIleGlnGlu-----AlaAspMetIleLeuProIleVal 323  
Db 1379 ATCTGCCAAGATCTGCAGAACCCCTCAGCTGTTCGACATCCGCTATCCCAATC 1438  
Qy 324 LeuGlnTyrLeuCysProValTyrIleSerPheGlyLeuGlyAlaValSerAlaIa 343  
Db 1439 GTGCTGGAATCTCTGCTCCACAGGCTCCGCTGATGATGCTGATGCTGATGCTGCT 1498  
Qy 344 ValMetSerSerAlaAspSerIleLeuSerAlaSerSerMetPheAlaIaGlnIle 363  
Db 1499 CTCATGTCTCTCCCTCACTCCATCTTAAACAGGACAGCACCTTCATCCATGACCTC 1558  
Qy 364 TyrGlnLeuSerPheArgIleAsnAlaSerAspIleGlnIleValTyrPheValMetArgIle 383  
Db 1559 TGGAAAT-----CACTCCGCGCTCGGGCATCTGAGAAGAGCTCATGATGTGGCAGGGTG 1615  
Qy 384 ThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIleValTyr 403  
Db 1616 TTGTG-----CTGCTGCTGGTCTGCTGCTCC 1642  
Qy 404 GlyLeuTyr-----TyrLeuSer 409  
Db 1643 ATCTCTGGATCTCCCTGTGCTCAGGCCAGCCAGGGCGGCACTTTCATCTATATCCAG 1702  
Qy 410 SerAspLeuValTyrIleIleIlePheProGlnLeuCysValLeuPheIleIleGly 429  
Db 1703 TCCATCAGCTCTACCTG-----CAGCCGCTGTGCGGGGTGCTTTCATCATGAGGA 1753  
Qy 430 -----ThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPhe 444

Db 1754 TGTTCGTGAAGAGACCAATGAAAAAGGTGCTTCTGGGGCTGTATCTGGGCTGCTC 1813  
Qy 445 LeuArgIleThrGly---GlyGlnProTyrLeuTyrLeuGlnProLeuIlePheTyrPro 463  
Db 1814 CTGGGCTGTGTTAGGCTGTGCTGCTGACCTTATTTATGCTGACGCTGATGCCACAGCA 1873  
Qy 464 GlyTyrTyrSerAspIleAsnGlyIleTyrAsnGlnArgPheProPheIleThrLeuSer 483  
Db 1874 -----GATAGGGCGCCGCTCGGTGAAGAGATTCATCACTACTTCTCC 1921  
Qy 484 MetValThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIleTyrIlePheGln 503  
Db 1922 ATGATCTGCTCCACGGCTCACCTTCATCACTGCTCTCCACCGTGGCTGCTTC----- 1972  
Qy 504 SerGlyThrLeuProProIleLeuValArgAsnGlnAlaIleIleLeuAsnGlnIleVal 523  
Db 1973 -----ACAGAGCTACCTCTCCAGAGAGATGTCACGACCTG----- 2008  
Qy 524 GluAsnMetAspIleThrIleLeuValArgAsnGlnAlaIleIleLeuAsnGlnIleVal 543  
Db 2009 -----ACCTGTTTACTGCTGTCACGACCCGCTGTCCAGAAAGAAACAGCA 2053  
Qy 544 ProValIleProArgGlnSerLeuThrLeuSerSerThrPheThrAsnIleGlnIleVal 563  
Db 2054 CCACGACGACTCCCTGTGCTCTTACCTCTCTCAG----- 2089  
Qy 564 LeuAspValAspSerProGlnIleGlySerGlyTyr 575  
Db 2090 -----AACGGATGCCAGAGGCCAGCAGCAGC 2116

RESULT 6  
US-10-162-012-28  
Sequence 28. Application US/10162012  
Patent No. 6682597  
GENERAL INFORMATION:  
APPLICANT: Curtiss, Roy A. J.  
APPLICANT: Silos-Santiago, Immaculada  
APPLICANT: Gu, Wei  
TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS  
FILE REFERENCE: 10448-190001  
CURRENT APPLICATION NUMBER: US/10/162,012  
CURRENT FILING DATE: 2002-06-04  
PRIOR APPLICATION NUMBER: US 60/209,845  
PRIOR FILING DATE: 2000-06-06  
PRIOR APPLICATION NUMBER: US 09/875,321  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: PCT/US01/18340  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/209,257  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: US 09/875,423  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: PCT/US01/18398  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/209,238  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: US 09/875,363  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: PCT/US01/18247  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/227,068  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: US 09/928,530  
PRIOR FILING DATE: 2001-08-13  
PRIOR APPLICATION NUMBER: PCT/US01/25475  
PRIOR FILING DATE: 2001-08-15  
PRIOR APPLICATION NUMBER: US 60/226,770  
PRIOR FILING DATE: 2000-08-21  
PRIOR APPLICATION NUMBER: US 09/934,421  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: PCT/US01/26096  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 60/279,281

```

      / PRIOR FILING DATE: 2001-03-28
      / PRIOR APPLICATION NUMBER: US 10/109,029
      / PRIOR FILING DATE: 2002-03-28
      / PRIOR APPLICATION NUMBER: PCT/US02/09728
      / PRIOR FILING DATE: 2002-03-28
      / PRIOR APPLICATION NUMBER: US 60/290,288
      / PRIOR FILING DATE: 2001-05-11
      / PRIOR APPLICATION NUMBER: US (not assigned)
      / PRIOR FILING DATE: 2002-05-13
      / NUMBER OF SEQ ID NOS: 48
      / SOFTWARE: FastSeq for Windows Version 4.0
      / SEQ ID NO: 28
      / TYPE: DNA
      / LENGTH: 2028
      / ORGANISM: Homo sapiens
      / US-10-162-012-28

Alignment Scores:
Pred. No.:      2,47e-25      Length:      2028
Score:          310.00      Matches:      149
Percent Similarity: 40.2%      Conservative: 111
Best Local Similarity: 23.0%      Mismatches:  224
Query Match:      10.4%      Indels:      164
DB:                3      Gaps:          27

US-10-724-806-4 (1-580) x US-10-162-012-28 (1-2028)

QY      8 leuValAlaIlelleuPheThyrlEuLeuIlePheleuValGlyIleTrrPaIaIatrrp 27
DB      85 ATCGGGGTCGTCAGTTCGTAATTCCTTCTTCTGCTGGCTGTCGTCAGTCATGTCACAGTC 144
QY      28 lYstHrLysAsnSerGlyAsnProGluGluArgSerGluAlaIlelleValGlyArg 47
DB      145 AAGACCAAA-----AGAGACAC AGTGAAGGCTACTCTCGCTGAAG 188
QY      48 AsplleGlyleuLeuValGly-GlyPheThrMetThrIaThrrPaValGlyGly 67
DB      189 GAACATGGTGGTGGCCAGTGGGTGCATCTGTTGGCCAGCAATGTGGAAGTGAC 248
QY      67 rIleAsnGlyThrAlaGluAlaValTrGlyProGlyCyGlyLeuAlaTrrPaIaHtsal 87
DB      249 TTTCATGGCCCTGGCA-----GGGTACGGTGGCTGACGGGCATTTTC 290
QY      87 adroIleGlyYrSerleuSer-----LeuIleLeuGlyleuPheP 102
DB      291 TGTATCAGCTTATGAATTAATGGCTGTTTCTGTGCTGATGTTGGCCGTGATCTTCC 350
QY      102 eaIaLysPrometArGSerleGlyYrValTrPheMetLeuAspProPheLysGlnIleTy 122
DB      351 A-----CCATCTCATTCGTCGTGTAGGTACCAAGATGCCAATTACTACGG----- 399
QY      122 rGlyLysArgMetGlyGlyleuLeuPheIleProAlaLeuMetGlyGluMerPheTrrPaI 142
DB      400 ----AAGCCCTTCGGGGCATGCA--ATCCCATCATCTCGCTGATCTTAACCTATT 452
QY      142 aAlaIaIlePheSerAlaLeuClYAlaHrrIle-----SerValIlelleasp-- 158
DB      453 TATCTACATCTTCACCAAGATTCGGGTAAACAATGACATGACAGGTGCCATCTTCACAGA 512
QY      159 -----ValAsPValAsnIleSerValIleValSerAlaIleuIleAlaIleLeuTy 175
DB      513 GTCCTTCGACACCTCGATCTGATCTGCTGGCCATAGTGGGGCTACTGGCCATCATCTGTGATA 572
QY      175 rThrlEuValGlyGlyleuTySerValAlaTyThrAsPValValGlnleuPheCysII 195
DB      573 CACGGTGTGGTGGCTGGCTGGCTGTGATCTACAGAGATGGCCCTCAGACGGCTGATCAT 632
QY      195 ePheIleGlyleuTrrPrlSerValProPheAlaLeuSerHisrProAlaValThrrsrl 215
DB      633 GCTTATGAAGAGGCTACCTTGAATGAGGCTAAC-----AGTTTGGCCGGGTGGT----- 661
QY      215 eGlyPheThrrAlaValHtsaIaLysTyT----- 224

```

Db	682	GGGATGGAGAGACTGAAAGGAAGTACTCTTGGCCCTGGCTAGCAACCGAGTGAAGA	740
Oy	224	-----	224
Db	741	CAGCAGCTGGCGGCTGCCCGGAGAGATGCCTTCATATTTTCCGAGATCGCTGATC	800
Oy	225	-GlnSerProTrpLeuGlyThrIleGlnSerValGluVal-----TyrThrTrpLe	241
Db	801	TGATCTCCCGGGGCGGGGGGCTCTATTGGAAATGCATCCATCCCTCTGGTACTGTG	860
Oy	241	uASPasnPhelLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTyrPheGlnArgVa	261
Db	861	CACCGAT-----CAGGTGATTTGCCAGCGGAC	887
Oy	261	IleuSerSerSerSerAlaThrTyrAlaGln-----ValLeuSerPheLeuAl	277
Db	888	TCTGGCTGCCAAGACCTGCTCCAGGCCAAAGAGGAGGCTCTGTAGTGGCTGATACCTGA	947
Oy	277	aAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSe	297
Db	948	GGTGCTGCCCTCTTCATTAATGGTGTCCCT-----GGAGTGGCACCGCAT	995
Oy	297	rThrAspTrpAsnGlnThrAlaTyrGlyTyrProAspProLysThrIySgluGln----	315
Db	996	CCTCTTCCAGATCAAGTGGCGCTGTGA-----GATCCAGAGATCTGCCAAGATCTG	1049
Oy	316	-----AlaAspMetIleLeuProIleValLeuGlnTyrLeuCysPr	329
Db	1050	CAGCAACCCCTCAGAGCTGTTCGACATCGGTAATCCAACTGCTGTCGAGAACTCTGCC	1109
Oy	329	oValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerAlaAs	349
Db	1110	CACAGGGCTCGTGAGGTGATGATGGGTGATGGTGGCGGCTCTCATGTCCTCCCTCAC	1169
Oy	349	pSerSerIleuSerIleAspSerMetPheAlaArgAsnIleTyrGlnIleuSerPheAr	369
Db	1170	CTCATCTTTAACAGTGCACGACACATCTTCACATGAGACTCTTGAAAT--CACCTCCG	1226
Oy	369	gGlnAsnAlaSerAspLysGluIleValITripValMetArgIleThrValLeuValPheG	389
Db	1227	GCCTCGGCGCATCTGAGAAGGAGCTCATGATGTGGCAGGGTGTGTTGG-----	1275
Oy	389	YAlaSerAlaThrAlaMetAlaLeuLeuThrLysThrValTyrGlyLeuTrp-----	406
Db	1276	-----CTGCTGCTGGTCTGGTCTCCATCTCTGGATCCCTGT	1313
Oy	407	-----TyrLeuSerSerAspLeuValTyrIl	415
Db	1314	GGTCCAGGCCAGCCAGGGCGGCGGCACTTCCTATATATACAGTCCATCAGCTCTTACT	1373
Oy	415	eIleIlePheProGlnIleuLeuCysValLeuPheIleLysGly-----Th	430
Db	1374	G-----CAGCGGCTGTGGGGGCTGCTTCATATGAGGATTTTCTGGAAAGAGAC	1424
Oy	430	rAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly--	449
Db	1425	CAATGAAAAGGAGTCTTTCGGGGCTGATTCGGGGCTGCTCTGGGCTTGGTAGGCT	1484
Oy	450	-GlyGluProTyrIleuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspLy	469
Db	1485	GGTCTGGACCTTATTATTCGACGCTTCGATGGACAGCAGCA-----GATGA	1532
Oy	469	sAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetValThrSerPhePh	489
Db	1533	GGCGCGGCTCTGGTGAAGAGACATTCATCTTACTTCTCCATGATCTGTGCACGGT	1592
Oy	489	eThrAsnIleCysValSerTyrIleuAlaLysTyrLeuPheGluSerGlyThrLeuProPr	509
Db	1593	CACCTTATCTACTGTCTCCACCGTGAGCTGTTC-----ACAGAGCCACC	1637
Oy	509	oLysLeuAspValPheAspAlaValAlaArgHisSerGluGlnLysMetAspLysTh	529
Db	1638	CTCCAGAGAGATGGTCAAGCCACCTGG-----AC	1664



```
OY 261 lLeuSerSerSerAlaThrTyrAlaGln-----ValLeuSerPheLeuAl 277
|||:||||: ||: ||||:||||: ||: ||||:||||:
1065 TCTGCTGCTCAAGAACCTGTCATGCGCAAGAGAGAGTCTCTGATGCGTCACTACCTGAA 1124
OY 277 aAlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSe 297
1125 GGTGCTGCCCCCTTTCATATATGTTGTTCCCT-----GGATGCTAGCCGCAT 1172
OY 297 rThraSPTrpaSngInThraAlaTyrGlyTyrProAspProIlySthrlYsgInGlu----- 315
|||:||||: |||: ||||:||||: |||: ||||:||||:
1173 CCTCTTCCAGATCAAGTGGCTGTGCA-----GATCCAGAGATCTGCCAGAAATCTG 1226
OY 316 -----AlaAspMetIleuProIleValLeuGlnTyrLeuCysPr 329
1227 CAGCAACCCCTCAGGCTGTGGACATGCGCATGCCAAACCTGCGGAACTCTGCGC 1286
OY 329 oValTyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerAlaAs 349
1287 CACAGGGCTCCGTGGGCTGATGATGCTGTGATGTTGGCGGCTTCATATGTTCCCTCAC 1346
OY 349 pSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeuSerPheAr 369
1347 CTCATCTTTTAAACATGTCAGACCATCTTCACATGACCTCTGGAAAT---CACCTCG 1403
OY 369 gGlnAsnAlaSerAspIlySgInIleValTTrpValMetArgIleThrValLeuValPheG 389
1404 GCCTGGGCATCTGGAAGAGACTCATGATGTGGGCAAGGTTGTGTG----- 1452
OY 389 yAlaSerAlaThrAlaMetAlaLeuThrIlySthValTyrGlyLeuTrr----- 406
1453 -----CTGCTGTGCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1490
OY 407 -----TyrLeuSerSerAspLeuValTyrIl 415
1491 GATCCAGGCCAGCCAGGCGGCCAGCTCTTCATCTATATCCAGTCCATCAGCTCCACCT 1550
OY 415 eIleIlePheProGlnIleuLeuCysValLeuPheIleIlySgIly-----Th 430
|||:||||: |||: ||||:||||: |||: ||||:||||:
1551 G-----CAGCCGCTGTGGCGGTGCTCTTCATCATGAGATGTTCTCGAAGAGAC 1601
OY 430 rAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeuArgIleThrGly-- 449
1602 CAATGAAAAGGTGCTCTTCTGGGCTGATCTCGGCGCTGCTCTCGGCTTGTTAGGCT 1661
OY 450 -GlyGlnProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyrTyrSerAspI 469
1662 GGTCTGTGACTTATTTAGTGCAGCTGATGCGACGCCA-----GATGA 1709
OY 469 sAsnGlyIleTyrArgGlnArgPheProPheIlySthIleuSerMetValThrSerPheP 489
1710 GCGCCCGGTCTGTGGAAGAGCATTCACCTTACTTCTCCATGATCTGTCACAGGT 1769
OY 489 eThraSnIleCysValSerTyrLeuAlaIlySthIleuPheGlnSerGlyThrLeuPro 509
1770 CACCTCATCACTGCTCTCCACCGTGAAGCGCTTC-----ACAGAGCCACC 1814
OY 509 oLysLeuAspValPheAspAlaValAlaArgHisSerGlnGluAsnMetAspIlyTh 529
1815 CTCCAAAGAGATGTCAGCCACTG-----AC 1841
OY 529 rIleLeuValArgAngIluAsnIleLysLeuAsnGlnIleuAlaProValIlyProArg 549
1842 CTGGTTTACTGCTCAGACCCGCTGCTCCAGAGAACAAGACCAACGACGCTCCCTT 1901
OY 549 nSerLeuThrLeuSerSerThrPheThrAsnIlySgInAlaLeuAspValAspSerSe 569
1902 GTCTCTTACCTCTCAG-----AACGGGAT 1928
OY 569 rProGlnGlySerGlyThr 575
1929 GCCAAGGCCAGCAGCAGC 1947
Db RESULT 8
```

```
US-09-799-451-234
; Sequence 234, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Yang, Yonghong
; APPLICANT: Weinman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 234
; LENGTH: 2107
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2057)
; US-09-799-451-234

Alignment Scores:
Pred. No.: 5.87e-25 Length: 2107
Score: 307.00 Matches: 130
Percent Similarity: 39.2% Conservative: 85
Best Local Similarity: 23.7% Mismatches: 192
Query Match: 10.3% Indels: 142
DB: 3 Gaps: 22

US-10-724-806-4 (1-580) x US-09-799-451-234 (1-2107)
OY 11 lIleIleuPheTyrLeuLeuIlePheLeuValGlyIleTrrAlaIleTrrIlySthrlYs 30
|||:||||: |||: ||||:||||: |||: ||||:||||:
174 ATTGATCATCTATTTTCTGTGGTATGCTGTGGTGTGGGCGATGCTGAAGACC--- 230
OY 31 AsnSerGlyAsnProGlnGluArgSerGluAlaIleValGlyAlaArgAspIleGly 50
231 AACGAGGACT-----ATAGGAGGCTTCTTCCTCGCTCGTCTGATATGCGC 278
OY 231 AAACGAGTACT-----ATAGGAGGCTTCTTCCTCGCTCGTCTGATATGCGC 278
OY 51 LeuLeuValGlyIlyPheThrMetThrAlaThrTrrValGlyGlyTyrIleAsnGly 70
279 TGGTGGCCGATGGGCGCTCTCTTTCGACATGACATGCGGACGACACCACTATGTGGGG 338
OY 71 ThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTrrAlaHisAlaProIleGly 90
339 CTGGGT-----GGACAGAGCGAGCTTCAAGAGTGCACACCGCTACA 380
OY 91 TyrSer-----LeuSerLeuIleuGlyGlyLeuPhePheAlaIlySthPro 105
381 TTTGATAGGACTTCTCTCAGTATGTTGCTGATTTCTGGTGGATGTTTGTCCCTATAC 440
OY 106 MetArgSerIlySgIlyTyrValThrMetLeuAspProPheIlySgInIleTyrGlyIlySg 125
441 ATCAAGTGG---GGGTGATGACCATGCCGGAATATCTCAAG-----AAGCGG 485
OY 126 MetGlyGly-----LeuLeuPheIlePro 133
```

Db	486	TTTGTGGGGAGCGACCTCCAGCTCTACTCTCCATCTCTCCCTTATCTGTGTGGTT	545
Qy	134	AlAlenuewGlyGluMetPheTrpAlaAlaIlePheSerAlaLeuGlyAlaThrIle	153
Db	546	TTGTAAATTTGTGAGACATATTTGTCTGGAGCCCATATTC-----ATC	587
Qy	154	SerValIleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAlaIle	173
Db	588	AAGCTGGCCTTGGGGATTGGACCTTATCTGGGAAATCTTCATCTCTTGGCTATGCTCT	647
Qy	174	LeuTrpThrIleuValGlyGlyLeuTrpIleSerValAlaTrpThrAspValValGlnIleuPhe	193
Db	648	GTTTAAACACCACTGGGGGCTTGGCCCTGGGATTTTACACAGACACCTCCAGACCATC	707
Qy	194	CysIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaValThr	213
Db	708	ATCATGTGATTTGGCTCTTATATCTCATGGGGTTTGATTTAAAGAAATTGA-----	761
Qy	214	AspIleGlyPheThrAlaValHisAlaIleAspTrpGlnSerProTrpLeuGlyThrIleGlu	233
Db	762	-----GGTTATGAGAGCTTTACCGAGAAAGTACGGAATGCCACCCATCCGTAGTCGAG	815
Qy	234	-----SerValGluValIleThr-----TrpLeuAspAsnIleuLeu	246
Db	816	GGGAGCAACTTGACATACATGACGCCAGTTGCTTACACCTCGGGCGGACTCTTCCACATC	875
Qy	247	Leu-----MetLeuGlyGlyIleProTrp-----	254
Db	876	TTCCGAGATGCTGTGACTGGGACATTCATGAGCCAGGAATTATTTGGAATGCCCAT	935
Qy	255	-----GlnAlaTrpPheGlnArgValLeuSerSerSer	265
Db	936	ACAGCTTTGTGGTACTGGTGCACAATACAGGTCAATGTGCAGCGCTGCTGTGTGGCAAG	995
Qy	266	SerAlaThrTrpAlaGlnValLeuSerPheLeuAlaIlePhe-----	279
Db	996	GACATGTCTACGTGAAGCGCGCTTGATTTGTGTGCTTACGTGAAGCTGCTGCCCATG	1055
Qy	280	-----GlyCys	281
Db	1056	TTCTCATAGTGATCGCGGGATGATTCAGCCGATCTCTACACAGATATGTAGTACATGT	1115
Qy	282	LeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsn	301
Db	1116	GTGGTA-----CCTTCTGAATGCGTGAACACACTGTGC-----GTTGATTTGTGC	1160
Qy	302	GlnIleThrAlaTrpGlyTrpProAspProIleThrIleGluAlaAspMetIleIleuPro	321
Db	1161	TGCACCAACTACGCATATCCCC-----ACGATGGGTCTGGAA	1196
Qy	322	IleValLeuGlnTrpLeuCysProValTrpIleSerPhePheGlyLeuGlyAlaValSer	341
Db	1197	CTGATGCCCCAGAGACTCGAGGGCCTGATGCTTGG-----GTATATGCTG	1241
Qy	342	AlAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArg	361
Db	1242	GCCCTCTCAAGAGCTCCCTGACCTCCATCTTCAACACAGCGCAGACCCCTCTTCCCAT	1301
Qy	362	AsnIleTrpGlnLeuSerPheArgGlnAsnAlaSerAspIleGluIleValTrpValMet	381
Db	1302	GACCTCTAC-----ACCAAGATCGGAGAGACAGCTCGAGAAAAGAGCTCTGATAGCTGGA	1355
Qy	382	ArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrIle	401
Db	1359	CGGATATTTGTTCTTCTTATTACTGTTGTGACCATGTGTGGGTCCTCGTATCAAGTT	1418
Qy	402	ValTrpGly-----LeuTrpTrpIleuSerSerAspLeuValTrpIle-----Ile	416
Db	1419	TCTCAAAATGAGCACTATATCATTCACACAGATCAATTTCTAGTACTCTTGGGCTCCA	1476
Qy	417	IlePheProGlnLeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTrpGlyAlaVal	436

Db	1479	ATTGACGCTGCTTTGCTGCTCCATCTTCTGTAAGAGTCAATGAAACAGGAGCATTC	1538
QY	437	AlaGlyTyrIlePheGlyLeuPheLeu-----	445
Db	1539	TGGGCTTAATGGTTGACTTCGAATGGGCTTCATTGATATGAAACAGAGTTGCTTAT	1598
QY	446	-----ArgIlethrGlyGlyProTyr	453
Db	1599	GGAACAGGAGTGTGCTGGCTCCAGTAATGTGCCAAGATTATCTGTGAGTGCACATAT	1658
QY	454	LeuTyrLeuGlnProLeuIlePheTyr	462
Db	1659	CTGTACTTTTCCATGCTGCTCTTTT	1685
RESULT 9			
	US-09-540-236-273		
	/ Sequence 273, Application US/09540236		
	/ Patent No. 6673910		
	/ GENERAL INFORMATION:		
	/ APPLICANT: Gary L. Breton et al.		
	/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS		
	/ FILE REFERENCE: 2709.2005-001		
	/ CURRENT APPLICATION NUMBER: US/09/540,236		
	/ CURRENT FILING DATE: 2000-04-04		
	/ NUMBER OF SEQ ID NOS: 3840		
	/ SEQ ID NO 273		
	/ LENGTH: 1566		
	/ TYPE: DNA		
	/ ORGANISM: M.catarrhalis		
	US-09-540-236-273		
Alignment Scores:			
	Pred. No.: 3,99e-24	Length: 1566	
	Score: 298.00	Matches: 132	
	Percent Similarity: 41.8%	Conservative: 99	
	Best Local Similarity: 23.9%	Mismatches: 223	
	Query Match: 10.0%	Indels: 100	
	DB: 3	Gaps: 18	
US-10-724-806-4 (1-580) x US-09-540-236-273 (1-1566)			
QY	9	ValAlaIleIleLeuPheTyrIleLeuIlePheLeuValGlyIleTrrAlaAlaTrrLys	28
Db	97	ATTTGCTGGCTGCTATATTATTTGATGATTCGAATCGAATTATGCTTATTTAAA	156
QY	29	ThrLysAsnSerGlyAsnProGluIuArgSerGluAlaIleIleValGlyIuArgAsp	48
Db	157	CAAAAATGAC-----ATTGAAGATATATGCTTGGCGGACGCAAT	198
QY	49	IleGlyLeuLeuAlaGlyIuPheThrMetThrAlaThrTrrValGlyIuTyrIle	68
Db	199	CTAATCTCTGCTGAATCATATTATCTGCAGGCGCTCTGATATGTCAGGCTGGCTTTTG	258
QY	69	AsnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTrrAlaHisAlaPro	88
Db	259	CTGGGCTTGCACAGATACATGACCTTCAAGCGTGTAAATGTTGG-----ATGCA	312
QY	89	IleGlyTyrSerLeuSerLeuIleLeuGlyIuPhePheAlaLysProMetArg---	107
Db	313	CTGGGCTTGACCATTTGGTGTGCTAATTATCTATGTTGCAACGCGTTTGGTGT	372
QY	108	-----SerLysGlyTyrValThrMetLeuAspProPheLysGlnIleTyrGly	123
Db	373	TATACCGAGCTTGCAATATAGCTGTACTTTACCAAGATTATTTTCAAAATCGTTTCAT	432
QY	124	LysArgMetGlyIuLeuLeuPheIleProAlaLeuMetGlyIuMetPhe-----	140
Db	433	GATAAATACATCTTTTGGTATCATGTCGCGCGGTGATCATATTATTTTAAACGTT	492
QY	141	TrrAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp	160
Db	493	TATACAGCAGCAGCTTGGTGTGTGTGTAACCTGTTGAAGACTTGAATTTCA	552

```
OY 161 ValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuThrThrLeuValGlyGly 180
DB 553 TATACAGATGGGTTTGGGTACACAGACGGCTTGTGGCTTATACCTTGTGGCCGC 612
OY 181 LeuThrSerValAlaThrThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
DB 613 TTTTGGAGGTGTCTTAACTGACTGCTTGTCCAAAGCGCTATTACTGTGATCCATGCTA 672
OY 201 IleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
DB 673 ATT-----GTAACCTGTGGTACATTTGGCCAGATT 702
OY 221 HisAlaLysThrGlnSerProTrpLeuGlyThrIleGluSerValGlnValThrTrp 240
DB 703 GGTGTGTGTCTGACGCGATGGCATGTGCACACAGACAAAATACGAGGTCTTTAATGG 762
OY 241 LeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaThr----- 257
DB 763 ATGAATGGC-----GTGACTGTCTATGGGTGTATCTCTTGTATGGCATGGGGTTTGGC 816
OY 258 -----PheGlnArgValLeuSerSerSerAla 267
DB 817 TATTTTGGTACCCACACATTTATGACCTTTATGGCAATTCCTCACTCAATTAAGATGT 876
OY 268 ThrTyraIaGlnValLeuSer-----PheLeuAlaIaPheGlyCysLeuVal 283
DB 877 CCGACTGCCATGATGGATGGCATGGGCTGATGATTTTAACTAATTTGCTTGTGATG 936
OY 284 MetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThr 303
DB 937 GTTGGCTTGGCAGGAGATGCTTATGTGCA-----CGCACA 972
OY 304 AlaTyraIyTyProAspProLysThrLysGlnIuaIaAspMetIleLeuProIleVal 323
DB 973 GGTATAGAGCTTAAGATCTGAAC-----ATCTTTTGGATTT 1014
OY 324 LeuGlnTyLeuCysProValTyIleSerPhePheGlyValAlaValSerAlaIa 343
DB 1015 TCACAGAGCTGTTCATCCGCTGATTCAGATTTGTTGGCGCAATTTTACAGACA 1074
OY 344 ValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsnIle 363
DB 1075 ATTATAGATCAATCTCAAGCAATTTGCTGTGCTCTCAAGTTCACTGACACAGAGTATT 1134
OY 364 TyrgInLeuSerPheArgGlnAsnAlaSerAspLysGlnIleValTyraIaMetArgIle 383
DB 1135 TATAGATTATTTTGGATTAACAAGCTCAGAGGACCGCAAGTGTGTTTGGCCGTATT 1194
OY 384 ThrValIleValPheGlyAlaSerAlaThrAlaMetAla-LeuLeuThrLysThrValTy 403
DB 1195 TCGGTGTTTATGTCGCGATCACTTGTGATTTGTGGCGAGGCAACAGATTCAAGTA 1254
OY 403 rGlyLeuTrpTyPyr-----LeuSerSerAspLeuValTyIleIleIleIlePheProGlnLe 421
DB 1255 TTTAAATTGTGTCTCATGCTTGGCAGAGATTTGGGCGACGATTTGGTCTATTG-CTAAT 1313
OY 421 uLeuCyValLeuPheIleIleGlyThrAsnThrTyrgIyaIaValaIaGlyTyriIlePhe 441
DB 1314 TTTAAGCCGTATGTG--AAGCGTATGAACCGAATAGTGCTTTGGCGCTGATGATTGT 1370
OY 441 eGly-----LeuPheLeuArgIleThrGlyGlyIuProTyLeuTyLeuGlnBr 458
DB 1371 GGGTCCCTTACGCGATTAATTTGGTGTATGTGT----- 1407
OY 458 cLeuIlePheTyProGlyTyTySerAspLysAsnGlyIleTyraSn-----GlnAr 476
DB 1408 -----TTTGAGATTGGCGGTACAGCCACTAATGATTCATTATTCATTATTTACCGG 1460
OY 476 gPheProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTy 496
DB 1461 TTTTCTTT-----AGTTGTGTACACACATTTGAGTAAGTCT 1499
```

```
OY 496 rLeuAlaLysTyLeuPheGluSerGlyThrLeuProProLysLeuAspValPheAspAl 516
DB 1500 GATG-----ACAGACCCCGCGCTGTATTATTTATTTCAAA 1535
OY 516 aValValaIaArgHisSerGluGluAsnMetAspLys 528
DB 1536 ATTT-----GAGGATATGGAAGAAG 1554

RESULT 10
US-09-596-002-7/c
; Sequence 7, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Paterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 6703
; TYPE: DNA
; ORGANISM: M. catarrhalis
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 7
; PUBLICATION INFORMATION:
US-09-596-002-7

Alignment Scores:
Pred. No.: 4,426-23 Length: 6703
Score: 298.00 Matches: 132
Percent Similarity: 41.88 Conservative: 99
Best Local Similarity: 23.94 Mismatches: 223
Query Match: 10.04 Indels: 100
DB: 3 Gaps: 18

US-10-724-806-4 (1-580) x US-09-596-002-7 (1-6703)

OY 9 ValAlaIleIleLeuPheTyLeuLeuIlePheLeuValGlyIleTyraIaAlaTrpLys 28
DB 2431 ATTTGTTGGCTGTCTATTATTTATTTGATGATTCGCAATCGGAATTTATGCTTATTTAAA 2372
OY 29 ThrLysAsnSerGlyAsnProGluGluArgSerGluAlaIleIleValGlyArgAsp 48
DB 2371 CAAAAAAATGAC-----ATTGAAGATATATATCTTGGCGGACGCAAT 2330
OY 49 IleGlyLeuLeuValGlyPheThrMetThrAlaThrTrpValGlyGlyTyriIle 68
DB 2329 CTAACTCTGTGCTGTAACGATATCTGACGTGCTGTGATATCTGACGTGCTTTTG 2270
OY 69 AsnGlyThrAlaGlnAlaValTyrgIyProGlyCysGlyLeuAlaTrpAlaHisAlaPro 88
DB 2269 CTGGGCTTGACAGATATACATGACCTTCAGGCGGTGAAGTATTGG-----ATCCCA 2216
OY 89 IleGlyTyriSerLeuSerLeuIleLeuGlyGlyLeuPhePheAlaLysProMetArg--- 107
DB 2215 CTGGGTGTGACCATTTGGTGTGCTTGTCTAATTAATTCATTTGTGCACCGCGTTGCGTGT 2156
OY 108 -----SerLysGlyTyTyraIaThrMetLeuAspProPheLysGlnIleTyrgIy 123
DB 2155 TATACGAGCTTGGCAGATTAATGCTGTACTTTACCAAGTATTTTCAATGCTTTTCAT 2096
OY 124 LysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGluMetPhe----- 140
DB 2095 GATAAATCACATCTTTTGCATATCGTCGCGGGGTGTCATTATCTTATTTTTCACGTT 2036
OY 141 TrpAlaIaIaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAspValAsp 160
```



```

Db      2035 TATACAGACGACGCTGGTGGCTGGTAAAGCTTTGAAAGCTATTGAATTATCA 1976
Qy      161 ValaenileSerValleValSerAlaLeuIleAlaIleLeuTyThrLeuValGlyGly 180
Db      1975 TATACATGGGTTTGGGTAAACAGACGCGTGGTGGCTTATACCTTGTGGCGGC 1916
Qy      181 LeuTySerValAlaTyThrAspValValGlnLeuPheCysIlePheIleGlyLeuTrp 200
Db      1915 TTTTGGCGGTGCTATTAAGCTTGTCCAAAGGCGTATTATGCTGATCCGCACTCA 1856
Qy      201 IleserValProPheAlaLeuSerHisProAlaValThrAspIleGlyPheThrAlaVal 220
Db      1855 ATT-----GTGACTGTGATACATTGGCGGAGATT 1826
Qy      221 HisAlaIleTyThrGlnSerProTrpLeuGlyThrIleGlnSerValGlnValTyThrTrp 240
Db      1825 GGTGGTGTCTGAGGCGATGGCGATGGCCACACAGACAAATAACGAGTCTTTAATTGG 1766
Qy      241 LeuAspAsnPheLeuLeuLeuLeuGlyGlyIleProTrpGlnAlaTy----- 257
Db      1765 ATGAATGCG-----GTGACTGTGATGGTGTATCTTTGATGGCAGGGGTTTGGC 1712
Qy      258 -----PheGlnArgValLeuSerSerSerAla 267
Db      1711 TATTTTGGTCAGCACACATTAATTGACGTTTATGGCAATTCGCTCACTTAAGATGTT 1652
Qy      268 ThrTyAlaGlnValLeuSer-----PheLeuAlaAlaPheGlyCysLeuVal 283
Db      1651 CCGACTGCGCATGGTGTATGGCATGGCGATGATTTTAAGCTTAATGGTGGTGGTGAAG 1592
Qy      284 MetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThr 303
Db      1591 GTTGGCTTGGCAGGAGTCTTATGTGCA-----CCGACA 1556
Qy      304 AlaTyGlyTyProAspProLyThrLyGlnGlnAlaAspMetIleLeuProIleVal 323
Db      1555 GGTATAGACCTTAAGATCTGTAAC-----ATCTTTTGGTATTT 1514
Qy      324 LeuGlnTyThrLeuCysProValTyThrIleSerPhePheGlyLeuGlnValAlaSerAlaIle 343
Db      1513 TCACAGAGCTGTTTCATCCGCTGATTTAGAGATTTTGTGGCGGCAATTTTACAGCA 1454
Qy      344 ValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerPheAlaArgAsnIle 363
Db      1453 ATTAATGAGTACATCTCAAGCCAAATTCGTGGTGTCTCAAGTTCACCTGACACAGATAT 1394
Qy      364 TyGlnLeuSerPheArgGlnAsnAlaSerAspLyGlnIleValTrpValMetArgIle 383
Db      1393 TATAAGTTATTTTGGATTAACAAGCCTCAGAGCGACGCCAAGTGTATTGGCGCTATT 1334
Qy      384 ThrValLeuValPheGlyAlaSerAlaThrAlaMetAla-LeuLeuThrLyThrValTy 403
Db      1333 TCGGTGTTTGTTCGCGATCTATGGATTTATGTGGCAGCGCATTCAGCAAGTTCAGTA 1274
Qy      403 rGlyLeuTrpTyR-----LeuSerSerAspLeuValTyRleIleIlePheProGlnIle 421
Db      1273 TTAATTTGGTATTCACATGCTTGGCGAGATTTGGGCGACGATTTGCTCATTCG-CTAAT 1215
Qy      421 uLeuCysValLeuPheIleLyGlyThrAsnThrTyRgylAlaValAlaGlyTyRlePhe 441
Db      1214 TTTAAGCCTGATGTG--AAGGCTATGAACCAATATGGGTGGTGGCGATGATTTG 1158
Qy      441 eGly-----LeuPheLeuArgIleThrGlyGlyGlnProTyLeuTyRLeuGlnTr 458
Db      1157 GGTGCTTGGACGCTATTATTTGGGTGATGTGTGT----- 1121
Qy      458 oLeuIlePheTyRProGlyTyRtyrSerAspLyAsnGlyIleTyRAsn-----GlnAr 476
Db      1120 -----TTTGAATTTGGCGGTGACGACAGTAAATGATTCATTTATTCACCGG 1068
Qy      476 gPheProPheLyThrLeuSerMetValThrSerPhePheThrAsnIleCysValSerTy 496

```

```

Db      1067 TTTTGCCTTT-----AGTTGGTGACAAACCATTCAGTAAGTCT 1029
Qy      496 rLeuAlaIleTyThrLeuPheGlnSerGlyThrLeuProProLyLsLeuAspValPheAspAl 516
Db      1028 GATG-----ACACACCGCGCGCTGTTTATATTGTTCAAA 993
Qy      516 aValValAlaArgHisSerGlnGlnAsnMetAspLyS 528
Db      992 ATTT-----GAGGATATGGAAGAAG 974

RESULT 11
US-09-949-016-5012
; Sequence 5012, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5012
; LENGTH: 2450
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5012

Alignment Scores:
Pred. No.: 1,43e-23 Length: 2450
Score: 296.00 Matches: 148
Percent Similarity: 38.1% Conservative: 99
Best Local Similarity: 22.8% Mismatches: 222
Query Match: 9.9% Indels: 180
DB: Gaps: 29

US-10-724-806-4 (1-580) x US-09-949-016-5012 (1-2450)
Qy      11 IleIleLeuPheTyRLeuLeuIlePheLeuValGlyIleTrpAlaAlaTrpLyThrLyS 30
Db      104 ATCGTATCTACTCTGCTGTATGATGGCGCGCTGCGACTGTGGCTATGTTTCCACC--- 160
Qy      31 AsnSerGlyAsnProGlnGlnArgSerGlnAlaIleIleValGlyGlyArgAspIleGly 50
Db      161 AATCGTGGACCT-----GTGGAGGCTTCTCTGCGAGCGCGAAGTATGCTG 208
Qy      51 LeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyTyRleAsnGly 70
Db      209 TCGTGGCGCATTTGAGAGCCCTCTTGTCTAGTAACATTTGAAGTGGCACCTTGGGG 268
Qy      71 ThrAlaGlnAlaValTyRgylProGlyCysGlyLeuAlaTrpAlaHisAlaProIleGly 90
Db      269 CTGGCC-----GGGACTGGGCGAGCTTCAGGCAATCGCCATTTGAGGC 310
Qy      91 TyRser-----LeuSerLeuIleLeuGlyGlyLeuPhePheAlaLySPro 105
Db      311 TTTGAATGAATGCCCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 364
Qy      106 MetArgSerLyS---GlyTyRValThrMetLeuAspProPheLySglnIleTyRgLyS 124
Db      365 ATCTATATTAAAGCTGGGTGTGTGACCAATGACAGTACTGCTGAGG-----AAG 412
Qy      125 ArgMetGlyGly-----LeuLeuPheIle 132
Db      413 CGGTTTGGAGGCGCAGCGGATTCAGGCTTACCTTTCCCTTCTGCTGCTGCTTACATT 472

```

```

QY 133 ProAlaLeuMetGlyGluMetPheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThr 152
Db 473 TTCACCAAGATCTCGGACAGACATCTTCGGGGCCATATTC----- 514
QY 153 IleserValIleIleAspValAspValAsnIleSerValIleValSerAlaLeuIleAla 172
Db 515 ATCAATCTGGCCCTTGGCCCTGATCTGTATTTAGCCATCTTCTCTATTGGCAATCACT 574
QY 173 IleLeuTyThrLeuValGlyGlyLeuTySerValAlaTyThrAspValValGluLeu 192
Db 575 GCCCTTTACAAATTAACAGGGGGCCCTGGGGGGATTTACAGGACACCTTGCAGAG 634
QY 193 PheCyIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHisProAlaVal 212
Db 635 GTGATCATGCTGGTGGGGCTTTAACTCCGACTGGGTTTGTTCACGAGGTGGGA--- 691
QY 213 ThrAspIleGlyPheThrAlaValAlaHisAlaTySerGluSerProTrpLeuGlyThrIle 232
Db 692 -----GGCTATGAGCCCTTCATGGAAGATCATGAAAGCC---ATTCACACCATTA 739
QY 233 GluSer-----ValGluValTyThr---TrpLeuAspAsn 244
Db 740 GTGCTCATGCAACACCACTTTCAGAAATAATGCTACATCCAAAGGCCGACCTCTTC 799
QY 245 LeuLeuLeu-----MetLeuGlyGlyIleProTrp----- 254
Db 800 CACATCTTCGAGATCCCTCAGGGAGACCTCCAGGCTGGGTTCATCTTGGGATG 859
QY 255 -----GlnAlaTyPheGlnArgValLeuSer 263
Db 860 TCCATCCTTACCTTGGTGTGCTGCTGACAGATCAGGTATGTGTGACGGCTGCTCTTA 919
QY 264 SerSerSerAlaThrTyThrAlaGln-----ValLeuSerPheLeuAlaAlaPhe 279
Db 920 GCCAAGAAATATGCTCACGTGAAGGGTGGCTGATCTGTGGGGATCTTAAGCTGAG 979
QY 280 GlyCySLeuValMetAlaLeuProAlaIle----- 289
Db 980 CCCATGTTTCATCATGTGATGCCAGGAAGATCAGCCGATCTGTACACAGAAAAATT 1039
QY 290 -----CysIleGlyAlaIleGlyAlaSerThrAspTrpAsnGln 302
Db 1040 GCCTGTGCTGCTCCCTTCAGAAATGTGAGAAATTTGGGGATTCACAGGTTGGCTTACCAAC 1099
QY 303 ThrAlaTyGlyTyPheProAspProLysThrLysGluGluAlaAspMetIleLeuProIle 322
Db 1100 ATGCGCTAT-----CCAAACC 1114
QY 323 ValLeuGlnTyLeuCyProValTyIleSerPhePheGlyLeuGlyAlaValSerAla 342
Db 1115 TTAGTGTGGAGCTCATGCGCAATGATGCGAGGCTGATGCTATCAGTCACTGCTGCC 1174
QY 343 AlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSerMetPheAlaArgAsn 362
Db 1175 TCCCTCATAGAGTCTCCGACCTCCATCTTCACACAGGCGACACCTCTTCACCATGAGC 1234
QY 363 IleTyGlnLeuSerPheArgGlnAsnAlaSerAspLysGluIleValTrpValMetArg 382
Db 1235 ATCTAC---GCCAAGTCCGCAAGAGAGAGCATCTGAGAAAGACCTCATGATTGCCGGAAG 1291
QY 383 IleThrValLeuVal---PheGlyAlaSerAlaThrAlaMetAlaLeuLeuThrLysThr 401
Db 1292 TTGTTATCTGTGCTGATTGGCATCAGCATCGCTGGGGTCCCATTTGTGATCAGCA 1351
QY 402 ValTyArgLy-----LeuTrpTyLeuSerSerAspLeuValTyThrIle-----IleIle 417
Db 1352 CAAATGGGCACTCTTCGATTACATCCAGTCCATACACAGATTACTTGGGACCAACCATTT 1411
QY 418 PheProGlnLeuLeuCyValLeuPheIleLysGlyThrAsnThrTyGlyAlaValAla 437
Db 1412 GCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471
QY 438 GlyTyThrIlePheGlyLeuPheLeuArgIle----- 447

```

```

Db 1472 GGACTGATCTTAGGACTTCTGATTGGGATTTACGTATGATTACTGATTGCTTATGGA 1531
QY 448 ThrGly-----GlyGluProTyLeu 454
Db 1532 ACCGGAGACTGCATGAGACCCAGCACTGTCCACAGATTATCTGGGGGTGCACACTTG 1591
QY 455 TyLeuGlnProLeuIlePheTyPheProGlyTyTyTySerAspLysAsnGlyIleTyAsn 474
Db 1592 TACTTGGCATTAATCTCTTC----- 1612
QY 475 GlnArgPheProPheLysThrLeuSerMetValThrSerPhePheThrAsnIleCyVal 494
Db 1613 -----GCCATTTCTATCATCCATCGTGTGATC 1642
QY 495 SerTyLeuAlaTyTyLeuPheGluSerGlyThrLeuProProLysLeuAspValPhe 514
Db 1643 TCCCTCTCCACCAACCATTT-----CCGATGTCGATCTTAC 1681
QY 515 AspAlaVal-----ValAlaArgHisSerGluGluAsnMetAspLysThrIleLeuVal 532
Db 1682 CGTCTGTGTGGAGCTCGCAACAGCAAGAGAGCGTATTGAC-----CTGGATGCG 1735
QY 533 ArgAsnGluAsnIleLysLeuAsnGluLeuAlaProValLysProArgGlnSerLeuThr 552
Db 1736 GAAAGAGAACATCCAAAGAGC-----CTTAAGAGACCATTTAA 1777
QY 553 LeuSerSerThrPheThrAsnLysGlu 561
Db 1778 ATGAAACACAAAGTTCCGAGAGAGAA 1804

RESULT 12
US-09-252-991A-7814/C
; Sequence 7814, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rudenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7814
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1036)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7814

Alignment Scores:
Pred. No.: 2,39e-23 Length: 3342
Score: 296.00 Matches: 123
Percent Similarity: 40.2% Conservative: 96
Best Local Similarity: 22.6% Mismatches: 230
Query Match: 9.9% Indels: 96
DB: Gaps: 16

US-10-724-806-4 (1-580) x US-09-252-991A-7814 (1-342)
QY 9 ValAlaIleIleLeuPheTyLeuLeuIlePheLeuValGlyIleTrpAla---AlaTrp 27
Db 1638 ATGGCGCTCGACATATTGCTGTGATCTTACCGCGGGGCGATGATGCGCTGGGCTGG 1579
QY 28 -----LysThrLysAsnSerGlyAsnProGluGluArgSerGluAla 41

```

```

Db      1578 TACGGCATCGCGCGCCAGAACCCGCGAGCAGC----- 1546
Qy      42  ILeIleValIGlyIValArgAspIleGlyLeuValIGlyIlePhe-----Thrm 58
Db      1545 TACCTGTGTGCGCGCGCGCAGCTCGGC-----CCGCGCTTCTACCTGGAGACCATG 1495
Qy      59  ThrAlaThrValGlyGlyGlyTyriLeAsnGlyThrAlaGluAlaValTyriGlyPro 78
Db      1494 GCGCGCAGCTCTCGCGCGCGCGCTCACCATGCGACCGTGGCGCTCGGCTACGTCCAT 1435
Qy      79  GlyCySGlyLeuAlaThrAlaHisAlaProIleGlyTyriSerLeuSerLeuIleuGly 98
Db      1434 GGCATCTCGCGCTTCTCGCGTGTGCGCGCGAGTCGCG-----CTCGCGCATCTGCGCGCTC 1381
Qy      99  GlyLeuPhePheAlaAlaValPheMetArgSerIleGlyTyriValIleMetLeuAspProPhe 118
Db      1380 AGCTGTGTCTCGCGCAGCGCGCTGTGAAGCTGAAGATCTACACCGCTCACCGAGTCTCG 1321
Qy      119  LysGlnIleTyriGlyIValArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyIle 138
Db      1320 GAGCGTGTGTACACACCGCGCGCGCGCGCGCGCGCGCTGATCATGCTGTGTACGCA 1261
Qy      139  MetPheTyriAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db      1260 CTGATGATCGCGCGCGCTGACCATGCGCATGCGCGCGCGCTCATGCGAGCTCTGTGCGC 1201
Qy      159  ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleLeuTyriThrIleVal 178
Db      1200 CTGCGCTTGTGGGTGTGTGATGCGCGCGCGCGCGCGCGCGTGTGTGTCTACCATC 1141
Qy      179  GlyGlyLeuTyriSerValAlaTyriThrAspValAlaGlnLeuPheCysIlePheIleGly 198
Db      1140 GCGCGCATGTGTGTGTGCGTGAACCGTGAACCATGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
Qy      199  Leu---TyrIleSerValIleProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe 217
Db      1080 CTGCGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033
Qy      218  ThrAlaValHisAlaIleTyriGlnIleSerProTyriLeuGlyIleThrIleGlnSerValIle 237
Db      1032 GACGCGCTGTGAGCAAGCTGCGCGCGCGAGTATCTC---GACTTACCGCGAGTGTGCTGG 976
Qy      238  TyrThrTyriLeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTyriGlnAlaTyri 257
Db      975  GACACCATGTGTACCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 916
Qy      258  PheGlnArgValIleuSerSerSerSerAlaThrTyriAlaGlnValIleuSerPheIleuAla 277
Db      915  TGGCAGCGGGGTGTTCACCGCGCGCGCGAGACCGTGGCCAAAGGTCCGCGCTCCGCGCC 856
Qy      278  AlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSer 297
Db      855  GGTATCTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 796
Qy      298  ThrAspTyriAspGlnThrAlaTyriGlyTyriProAspProIleThrIleGlnGluAlaAsp 317
Db      795  CTG-----CTGCGCGGAC-----CTGCGAGAACGTCAAC 769
Qy      318  MetIleLeuProIleValLeuGlnTyriLeuCysProValTyriIleSerPheGlyLeu 337
Db      768  AACGCGCTTGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 709
Qy      338  GlyAlaValSerAlaValMetSerSerAlaAspSerSerIleuSerAlaSerSer 357
Db      708  GCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 649
Qy      358  MetPheAlaArgAsnIleTyri---GlnLeuSerPheArgGlnAsnAlaSerAspIleGln 376
Db      648  ACCGTACACCGAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 589
Qy      377  IleValTyriValMetArgIleThrValIleValPheGlyAlaSerAlaThrAlaMetAla 396
Db      588  GAGGTGTACGAGAACCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 529

```

```

Qy      397  LeuLeuThrIleThrValTyriGlyLeuTyriLeuSerSerAspLeuValTyriIleIle 416
Db      528  CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 469
Qy      417  IlePheProGlnLeuLeuCysValLeuPheIleIleGlyIleThrAsnThrTyriGlyAlaVal 436
Db      468  AGCTGATCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 409
Qy      437  AlaGlyTyriIlePheGlyIleuPheLeuArgIleThrGlyGlyIleProTyriLeu 456
Db      408  ACCAGCATGACCTGTGCG-----TTCTCT 385
Qy      457  GlnProIleIlePheTyri-----ProGlyTyriTyriSer 467
Db      384  ACCGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 325
Qy      467  ----- 467
Db      324  CTGCGCGTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 265
Qy      468  -----AspIleAsnGlyIleTyriAsnGlnAla 476
Db      264  GTGCGCAGCGCGCGCTGTGAACGAGCAACCGCGAGCGAGGAGCTTCTCTGCTGCTT 205
Qy      477  PheProPheTyriThrIleuSerMetValThrSerPheThrAsnIleCysValSerTyri 496
Db      204  TTCATTTCCCGAGACATGAAAAAGAGACGGCATGACCAAC-----TGGCGGAATTC 151
Qy      497  LeuAlaIleTyriLeu 501
Db      150  CTGTGTACGACCACTC 136

```

## RESULT 13

```

US-09-252-991A-7678
; Sequence 7678, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7678
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (604)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7678

```

## Alignment Scores:

```

Pred. No.: 1.18e-23 Length: 1455
Score: 293.50 Matches: 114
Percent Similarity: 42.5% Conservative: 90
Best Local Similarity: 23.8% Mismatches: 213
Query Match: 9.8% Indels: 63
DB: 3 Gaps: 14

```

US-10-724-806-4 (1-580) x US-09-252-991A-7678 (1-1455)

```

Qy      9  ValAlaIleIleLeuPheTyriLeuIlePheLeuValIGlyIleTyriAla---AlaTyri 27
Db      2  ATGGCGCTGTGACATATTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 61

```

```
QY 28 -----LysThrIysAsnSerGlyAsnProGluGluArgSerGluA 41
Db 62 TACGGCATGCGCGCGCAAGACCCGCGACGAC----- 94
QY 42 IleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPhe-----ThrMet 58
Db 95 TACCTGGGCGCGCGCGCACTCGGC-----CCGGGCTTCTACCTGGGAACCATG 145
QY 59 ThrAlaThrTrpValGlyGlyGlyTrpIleAsnGlyThrAlaGluAvalAlaValTrpGlyPro 78
Db 146 GCCGCGACCGCTCTGGCGCGCGCTCCACCATCGGACCGTGGCGCTGGCTACCGTCAT 205
QY 79 GlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTrpSerLeuSerLeuIleLeuGly 98
Db 206 GGCATCTCCGGCTTGGCTGGCGCGCGATCGGC-----CTCGGCATCTCGTCCGCCCTC 259
QY 99 GlyLeuPhePheAlaIlePheMetArgSerLeuGlyTrpValThrMetLeuAspProPhe 118
Db 260 AGCTGTTCTGGCGCGCGCGCTGCTGAAGCTGAAGATCTACACCGTCCACCGAGTCTCG 319
QY 119 LysGlnIleTrpGlyLysArgMetGlyGlyLeuLeuPheIleProAlaLeuMetGlyGlu 138
Db 320 GAGCGTCCGTACAACCCGCGCGCGCGCACCGCGCGCGCTGATCATGCTGGTCTACGCA 379
QY 139 MetPheTrpAlaAlaAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleIleAsp 158
Db 380 CTGATGATCGGCGCGCACCTCGACCATCGCATCGGACCGCTACAGCTCTGTTCCGC 439
QY 159 ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleThrTrpThrLeuVal 178
Db 440 CTGCGCTTCTGGGCTTCATCTGATCGCGCGCGCGCTGGTGTCTTACTTCCACCATC 499
QY 179 GlyGlyLeuTrpSerValAlaTrpThrAspValAlaGlnLeuPheCysIlePheIleGly 198
Db 500 GCGCGCATGTTGGTCCCTGACCTGACCGCATCGTGCAGTTCTGATCATGACCTCGGC 559
QY 199 Leu---TrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe 217
Db 560 CTGCTGTTCTGCTGATCGCGCTGTCGATCAACGACCGCGC-----GAGTGG 607
QY 218 ThrAlaValHisAlaIleLysTrpGlnSerProTrpLeuGlyThrIleGlnSerValGluVal 237
Db 608 GACGCGCTCGTGAAGCCAGCTGCGCGCGCACGTACTTC---GACTTCAACCGCATCGGCTGG 664
QY 238 TyrThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGlyGlyIleProTrpGlnAlaTrp 257
Db 665 GACACCATCGTCACCTACTTCTGATCTTCTTCCGCGATCTTCATCGCGCACGACATC 724
QY 258 PheGlnArgValAlaLeuSerSerSerAlaThrTrpAlaGlnValLeuSerPheLeuAla 277
Db 725 TGGCAGCGGAGTTCACCGCGCGCGACGACCGTGGCCAGTGGCGCGCTCCGCGCC 784
QY 278 AlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSer 297
Db 785 GGTATCTACTGCTGCTCTACGCGATGGCGGGGCGTGAATGGAGATGGCGCGCAAGCTC 844
QY 298 ThrAspTrpAsnGlnThrAlaTrpGlyTrpProAspProLysThrLysGluGluAlaAsp 317
Db 845 CTG-----CTGCGCGAC-----CTGCGAAGAACGTCAAC 871
QY 318 MetIleLeuProIleValLeuGlnTrpLeuCysProValTrpIleSerPheGlyLeu 337
Db 872 AACGCTTTCGCCAGTGGTGCAGACAGCTGCGCAACGCGCATCGCGCGCTGGTCACTC 931
QY 338 GlyAlaValSerAlaAlaValMetSerSerAlaAspSerSerIleLeuSerAlaSerSer 357
Db 922 GCCGCGCGCGCGCGCGCTGATGTCCACCGCGACGCGCGCTCTGGCGCGCTCCACC 991
QY 358 MetPheAlaArgAsnIleTrp---GlnLeuSerPheArgGlnAsnAlaSerLysGlu 376
Db 992 ACCGTACCCAGGACCTGCTGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1051
QY 377 IleValTrpValMetArgIleThrValLeuValPheGlyAlaSerAlaThrAlaMetAla 396
```

```
Db 1052 GACGGACGAGAAACCGCATCGCGACCTGCTGGCGCTGATAGTCTCGCATTCGC 1111
QY 397 LeuLeuThrLysThrValTrpGlyLeuTrpTrpLeuSerAspLeuValTrpIleIle 416
Db 1112 CTGGTGTTCAGCGACGTATATACGCCCTTGACCGTGGCTTACACCTCTGTTAGCGGC 1171
QY 417 IlePheProGlnLeuLeuCysValLeuPheIleLysGlyThrAsnThrTrpGlyAlaVal 436
Db 1172 ATGCTGATCCCGCGATGATGGCGCGCATCTGAAGCGCGCACCGCGCGCGCGCATC 1231
QY 437 AlaGlyTrpIlePheGlyLeuPheLeuArgIleThrGlyGlyGluProTyLeuTrpLeu 456
Db 1232 ACCAGCATGACCTTGGGC-----ProGlyTrpTrpSer 467
QY 457 GlnProLeuIlePheTrp-----ProGlyTrpTrpSer 467
Db 1256 ACCGTACTGGTGTTCATATCAAGACGCGCTCCAGCCCATACCGCATCTACTACAGC 1315

RESULT 14
US-09-252-991A-7528
; Sequence 7528, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubinfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7528
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (696)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-7528

Alignment Scores:
Pred. No.: 1,22e-23 Length: 1485
Score: 293.50 Matches: 114
Percent Similarity: 42.5% Conservative: 90
Best Local Similarity: 23.8% Mismatches: 213
Query Match: 9.8% Indels: 63
Gaps: 14

US-10-724-806-4 (1-580) x US-09-252-991A-7528 (1-1485)
QY 9 ValAlaIleIleLeuPheTrpLeuLeuIlePheLeuValGlyIleTrpAla---AlaTrp 27
Db 94 ATGCGCTTCGACATATTCGCTGCTGATCTTACCGCGCGGAGATGATGCGCTGGGCTGG 153
QY 28 -----LysThrIysAsnSerGlyAsnProGluGluArgSerGluA 41
Db 154 TACGGCATGCGCGCGCGCAAGACCCGCGACGAC----- 186
QY 42 IleIleValGlyGlyArgAspIleGlyLeuValGlyGlyPhe-----ThrMet 58
Db 187 TACCTGGGCGCGCGCGCACTCGGC-----CCGGGCTTCTACCTGGGAACCATG 237
QY 59 ThrAlaThrTrpValGlyGlyGlyTrpIleAsnGlyThrAlaGluAvalAlaValTrpGlyPro 78
Db 238 GCCGCGACCGCTCTGGCGCGCGCTCCACCATCGGACCGTGGCGCTGGCTACCGTCAT 297
QY 79 GlyCysGlyLeuAlaTrpAlaHisAlaProIleGlyTrpSerLeuSerLeuIleLeuGly 98
```

Db	298	GGCATCTCCGGCTTCTGGCTGCGGCGCATCGAC-----CTCGGCATGTCGGCGCTC	351
Qy	99	GlyLeuPhePheAlaIysPheMetArgSerIysGlyTyrValThrMetLeuAspProPhe	118
Db	352	AGCTGTCTCCGGCCAAAGCCGCTCTGAAGCTGAAGATCTACCGTCGACCCAGGTCCTG	411
Qy	119	LysGlnIleTyrGlyTylsArgMetGlyIleuLeuPheIleProAlaLeuMetGlyIu	138
Db	412	GAGGCTGGCTACAAACCCGCGCGCGCCAGCCAGCGCCGATCATGCTGTCTTACGCA	471
Qy	139	MetPheTrpAlaIleAlaIlePheSerAlaLeuGlyAlaThrIleSerValIleLeasp	158
Db	472	CTGATGATCGGCGCACCTCGACCATCGCCATCGGACCCGTCAGTCAGAGTCTCTTCCGC	531
Qy	159	ValAspValAsnIleSerValIleValSerAlaLeuIleAlaIleTyrIleThrVal	178
Db	532	CTGCGCTTCTGGGTCTGCATCTCTGATCGGCGCGCGCTCGTGTCTTCTTCCACCATC	591
Qy	179	GlyIleuTyrSerValAlaIaTyrThrAspValAlaGlnLeuPheCysIlePheIleGly	198
Db	592	GGCGGATGATGTGTGCTACCCCTGACCGACATCGGAGCTTCGATCATGACCGCGGC	651
Qy	199	Leu---TrpIleSerValProPheAlaLeuSerHisProAlaValThrAspIleGlyPhe	217
Db	652	CTGGTGTCTCCGTGATCGCGCTCGCTGCATCAACAGACGCCGC-----GGATGG	699
Qy	218	ThrAlaValHisAlaIaTyrGlnSerProTrpLeuGlyThrIleGlnSerValGluVal	237
Db	700	GACGCCCTCGGTAGCCAGCTGCCGCGCACGACTTC---GACTTCACCGCATGGCGTGG	756
Qy	238	TyrThrTrpLeuAspAsnPheLeuLeuLeuMetLeuGlyIleProTrpGlnAlaTyr	257
Db	757	GACACCACTCCGACCTACTCTCTGATCTACTCTTCGCGACATCTTCATCGCCAGACATC	816
Qy	258	PheGlnArgValIleLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAla	277
Db	817	TGGCAGCGGGGTTCACCGCCCGCAGCAGACCGTGGCCAAAGTGC CGGCTCCGCGGC	876
Qy	278	AlaPheGlyCysLeuValMetAlaLeuProAlaIleCysIleGlyAlaIleGlyAlaSer	297
Db	877	GGATCTACTCGTGCATCGATGAGCGGGGGCGTGCATCGGATGAGCCCAAGATC	936
Qy	298	ThrAspTrpAsnGlnThrAlaTyrGlyTyrTrpAspProIysTrpGlnGluAlaAsp	317
Db	937	CTG-----CTGCCGAG-----CTGGACCAAGCTCAAC	963
Qy	318	MetIleLeuProIleValLeuGlnIleTyrLeuCysProValTyrIleSerPhePheGlyLeu	337
Db	964	AACGCTTCGGCAGCTGTGTCGAGACAGACCTCGCGAAGCGCATCGCGGCTGTATC	1023
Qy	338	GlyAlaValSerAlaIleAlaIleValMetSerSerAlaAspSerSerIleLeuSerAlaSerSer	357
Db	1024	GCCGCGCGCGCTGGGGCGCTGATGTCCACCGCCAGCGCGGCTGTGGCGGCTCCACC	1083
Qy	358	MetPheAlaArgAsnIleTyr---GlnLeuSerPheArgGlnAsnAlaSerAspIysIu	376
Db	1084	ACCGTCACCCAGACCTCTGCGCGGCTTGGCGCGCGCGCGCGCGGCGGCAAGTCCGACACGCG	1143
Qy	377	IleValIleTrpValMetArgIleThrValLeuValPheGlyAlaSerAlaIleThrAlaMetAla	396
Db	1144	GACGTCAGCAGAAACCGCATCGCGACCTCTGCTGGCTGTGTATGTCCTCGCATCGCC	1203
Qy	397	LeuLeuThrIysTrpIleValTyrGlyIleuTrpTyrLeuSerSerAspLeuValTyrIleIle	416
Db	1204	CTGGGTGTCAGCGACGATGATACGCGCCCTGACCGGCGGCTTCAACATCTGTGTAGCGCGC	1263
Qy	417	IlePheProGlnLeuLeuCysValLeuPheIleIysGlyIleThrAsnThrTyrGlyAlaVal	436
Db	1264	ATGCTGATCCGCTATATCGGGCGCATCTTACGTGAAGCGCGCACACCGCGCGCGCATTC	1323
Qy	437	AlaGlyTyrIlePheGlyLeuPheLeuArgIleThrIleGlyIleGluProTyrLeuTyrIleu	456
Db	1324	ACCAAGATGACCTCGGC-----TTCCTT	1347

```

QY      457 GlnProLeuIlePheTy-----ProGlyTyrIleSer 467
DB      1348 ACCGACTGGCTTCATCATCAGACGGCGCTCCAGCGCAATACGCCGATCTACTACACGC 1407

RESULT 15
; Sequence 1, Application US/07841651
; Patent No. 5410031
GENERAL INFORMATION:
APPLICANT: Pajor, Ana M
APPLICANT: Wright, Ernest M
TITLE OF INVENTION: Cloning and Functional Expression of a
TITLE OF INVENTION: Mammalian Na+/Nucleoside Cotransporter: A Member of the
TITLE OF INVENTION: SGLT Family
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Sheldon & Mak
STREET: 225 South Lake Avenue, Ninth Floor
CITY: Pasadena
STATE: California
COUNTRY: USA
ZIP: 91101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/841,651
FILING DATE: 19920224
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mandel, Saralynn
REGISTRATION NUMBER: 31,853
REFERENCE/DOCKET NUMBER: 8772
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2238 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Oryctolagus cuniculus
FEATURE:
NAME/KEY: CDS
LOCATION: 7..2022
US-07-841-651-1

Alignment Scores:
Pred. No.:          3,05e-22           Length:         2238
Score:              284.00             Matches:         150
Percent Similarity: 36.2%              Conservative:    82
Best Local Similarity: 23.4%            Mismatches:     226
Query Match:        9.5%               Indels:         182
DB:                 2                  Gaps:           26

US-10-724-806-4 (1-580) x US-07-841-651-1 (1-2238)

QY      8 LeuValAlaIleIleLeuPheTyrLeuLeuIlePheLeuValGlyLettPalaIaATTP 27
DB      82 ATCGGGCATTTGCTGTATTCTCTGGTCATTTGGTGCGCTTGCTGTCATGTCGC 141
QY      28 LysThrIlyAsnSerGlyAsnProGluGluArgSerGluAlaIleIlvalaIglyIarg 47
DB      142 AGAAC---AACAGAGCAC--------GTGGGTGGCTACTTCTCTGCAGACGA 186

```

```

QY 48 AsplleGlyLeuLeuValGlyGlyPheThrMetThrAlaThrTrpValGlyGlyTyr 67
   :::::
Db 187 AGCATGTGTGTGGCCGGCTGCTCTCTCTGACAAATATCGGAGTGGCCAC 246
QY 68 IleasnGlyThrAlaGluAlaValTyrGlyProGlyCysGlyLeuAlaTrpAlaHisIle 87
   |||||
Db 247 TTTGTGGGCGCTGGCC-----GGGACCGGTGTGCAACCGCTTGGCT 288
QY 88 ProIleGlyTyrSerLeuSer-----LeuIleLeuGlyGlyLeuPhePheAla 103
   :::::
Db 289 GTGGCGGATTTGAATGGAATGCGCTGTTCGTGGTGGCTCTCGGGTTGGCTGTTCGG 348
QY 104 LysProMetAlaGlySerGlyGlyTyrValThrMetLeuAspProPheGlyGlyIleTyrGly 123
   :::::
Db 349 CCGGTGTAACTCGACCGACGCGCTCATTCAGATG-----CCGAGTACCTGGCC 396
QY 124 LysArgMetGlyGly-----LeuLeuPhe 131
   |||||
Db 397 AAGCGCTTGGGCGGACATCGATCCGCTCTACTCTGCTCGCTTCTTCTGTATC 456
QY 132 IleProAlaLeuMetGlyGlyMetPheTrpAlaAlaAlaIlePhe-----SerAlaLeu 149
   :::::
Db 457 ATCTTCACCAAGATCTCGGTGACATGTTCGCGGGGCGGTGTATTACAGACGCTCTA 516
QY 150 GlyAlaThrIleSerValIleIleAspValAspValAsnIleSerValIleValSerAla 169
   |||||
Db 517 GGCTGGAAATATTACGCT-----TCGTCATTCGCGCTCTCTG 552
QY 170 LeuIleAlaIleLeuTyrThrLeuValGlyGlyLeuTyrSerValAlaTyrThrAspVal 189
   |||||
Db 553 GGCATCACCACATGCTTACACCTGACAGAGGCTGGACGCGCTATGTACACACACACA 612
QY 190 ValGlnLeuPheCysIlePheIleGlyLeuTrpIleSerValProPheAlaLeuSerHis 209
   |||||
Db 613 GTGCAGACCTTTGTCATATCGCGGGGCGCTTCATCTCACCGGTACGCGCTTCCACAG 672
QY 210 ProAlaValThrAspIleGlyPheThrAlaValHisAlaIleTyr----- 224
   |||||
Db 673 GTGGGC-----GGGTATTCGCGGCTCTTCGACAAATACATGGAGCGATGACT 720
QY 225 -----GlnSerProTrpLeuGlyThrIleGlnLeuSerValGluValTyrThr 239
   :::::
Db 721 TCGCTGACGCTGTCCGAGAGCCGCGCTGTGGCAACATCTCCAGCTCTCTGACCGACCC 780
QY 240 TrpLeuAspAsnPheLeuLeu-----MetLeuGlyGlyIleProTrp----- 254
   |||||
Db 781 CGGCTGACTCTATCATCTGCTCCGGACCTGTGACGGGGGACCTACCATGCGCCGG 840
QY 255 ----- 258
   |||||
Db 841 CTGCTCTGGGGCTCACCATGCTCTCGGGCTGTGACTGTGACGTGACCATGATAGTA 900
QY 259 GlnArgValLeuSerSerSerSerAlaThrTyrAlaGlnValLeuSerPheLeuAlaIle 278
   |||||
Db 901 CAGCGCTGCTGCGCGGAGAGAACCTGACCACTCAAGAGCA----- 942
QY 279 PheGlyCysLeuVal----- 283
   |||||
Db 943 ---GGCTGCATCTTGTGTGCTACTGAAGTACGCCCATGTCTCATGTGATGCCA 999
QY 284 -----MetAlaLeuProAlaIleCys 290
   |||||
Db 1000 GGAATGATCAGCCGATCCTTTACCCCTGACGAGGTGGCGTGGCGCTGAGGTGTGT 1059
QY 291 IleGlyAlaIleGlyAlaSerThrAspTrpAsnGlnThrAlaTyrGlyTyrProAspPro 310
   |||||
Db 1060 AAGCGCGGTGTGGACGGAAGTGGCTGTCCAAACATCGCTAT----- 1104
QY 311 LysThrIysGluGluAlaAspMetIleLeuProIleValLeuGlnTyrLeuCysProVal 330
   |||||
Db 1105 -----CCGCGCTGCTGTGTGAGTCTCATGCGCCAAAC 1134
QY 331 TyrIleSerPhePheGlyLeuGlyAlaValSerAlaAlaValMetSerSerIleAspSer 350

```

```

Db 1135 GATCGCCGCGACTCATCTGCGGCTCATGTGGCCGCGCTCATGCTCTCCCTGGCCCTCC 1194
   :::::
QY 351 SerIleLeuSerAlaSerSerMetPheAlaArgAsnIleTyrGlnLeuSerPheArgGln 370
   |||||
Db 1195 ATCTTCAACAGACAGACAGACTCTCTTACCATGACATCTACAGCTG-----CGGCC 1248
QY 371 AsnAlaSerAspLysGlyIleValAlaTrpValMetArgIleThrValLeuValPheGlyAla 390
   |||||
Db 1249 CGCGCCGCGCAAGCGAGAGCTCTGCTAGTACGAGCGCTCTGGGTGTTCATCGTGCG 1308
QY 391 SerAlaThrAlaMetAlaLeuLeuThrIysThrValTyrGly-----LeuTrpTyr 407
   :::::
Db 1309 GTGTGCGGTGGCTGCTACCTGTGTGACGGGACAGCGGCGGCGACCTTTCGATTAC 1368
QY 408 LeuSerSerAspLeuValTyrIleIle-----IlePheProGlnLeuLeuCysValLeu 425
   :::::
Db 1369 ATCCAGTCCGTTTCCAGTACTTGGCCCCCGCTGTGTGCAATCTTGTGTGGCGCTC 1428
QY 426 PheIleIysGlyThrAsnThrTyrGlyAlaValAlaGlyTyrIlePheGlyLeuPheLeu 445
   |||||
Db 1429 TTGCTGCGCGCGTTAATGAGAAGGCGCTTGTGGGAGCTGATAGAGGGGCGCTGCTATG 1488
QY 446 ArgIleThrGlyGlyGluProTyrLeuTyrLeuGlnProLeuIlePheTyrProGlyTyr 465
   |||||
Db 1489 GGCCTGGACGCGCTTATTCCTC----- 1509
QY 466 TyrSerAspLysAsnGlyIleTyrAsnGlnArgPheProPheLysThrLeuSerMetVal 485
   |||||
Db 1510 -----GAGTTCTCTTCCTGGGACCGGACGCTGCGTG 1539
QY 486 -----ThrSerPhePheThrAsnIleCysValSerTyrLeuAlaIysTyr 500
   :::::
Db 1540 CGACCTCTGCTTCCCGCGGACTTCTGTGTGGGACACTACTTACTTGGCCATGTG 1599
QY 501 LeuPheGlu-----Sergly-----ThrLeuPro---Pro 509
   |||||
Db 1600 CTCTTCTTCTGCTCTGGCTCTTCATCATCATCTCTCTGTCGACGTGACCATGCCA 1659
QY 510 LysLeuAspValPheAspAlaValAlaIle---ArgHisSerGluGluAsnMetAspLys 528
   :::::
Db 1660 CGCAAGACCTTCCACCGCTGTGTTCAGTCTCCGGGACAGACAGAG----- 1707
QY 529 ThrIleLeuValArgAsnGluAsnIleLysLeuAsnGluLeu---AlaProValIysPro 547
   |||||
Db 1708 -----GAAAGGGAAGACCTGATCTGACAGGCTGGAAGCCCGGCTCTCC 1755

```

Search completed: July 10, 2006, 14:15:20  
 Job time : 340 secs